

CUERPOS DE ESTADÍSTICOS DEL ESTADO

OPOSICIONES

EXÁMENES RESUELTOS



JULIO CÉSAR HERNÁNDEZ SÁNCHEZ



MIGUEL GARCÍA DÍAZ



**CUERPO DE
DIPLOMADOS**

**CUERPO
SUPERIOR**

PARTE COMÚN
ESPECIALIDAD CIENCIA DE DATOS

OPOSICIONES A LOS CUERPOS DE ESTADÍSTICA DEL ESTADO
EXÁMENES PRÁCTICOS RESUELTOS

© Publicaciones Instituto Nacional de Estadística, Madrid 2025

Julio César Hernández Sánchez

Miguel García Díaz

Edición: 1^a

Nº de páginas: 155

Actualizado a 24 de Julio de 2025

Ilustración de la portada y diseño gráfico:

Tomás Hernández Sánchez

*A Isa, Julia y Teresa, que son mi vida.
A mi hermano Tomy, un genio.
A mis padres que me enseñaron la humildad.*

*A mis padres, mi hermana y Mario por estar ahí siempre.
A Silvia, Bego y Dulce por todo lo compartido, la luz y el optimismo.
Y a esos buenos profesores.*



Julio César Hernández Sánchez

Licenciado en Matemáticas
Ingeniero Técnico en Informática de Sistemas
Estadístico Superior del Estado desde 2005
Delegado Provincial del INE en Zamora



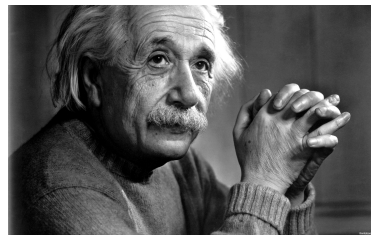
Miguel García Díaz

Licenciado en Matemáticas
(Ciencias de la Computación)
Máster en Finanzas Cuantitativas (AFI)

*Mathematics is the most beautiful and
most powerful creation of the human spirit.*
Stefan Banach

*I have had my results for a long time: but I do not yet know
how I am to arrive at them.*
Carl Friedrich Gauss

Agradecimientos



Try not to become a man of success but a man of value.

Albert Einstein

Nos gustaría agradecer el esfuerzo de todas las personas que han hecho posible que este documento pueda ver la luz y pueda ayudar a aquellos opositores ilusionados con este camino. En particular queremos hacer una mención especial a Cristobal Rojas Montoya, María de Blas Portero, Román Mínguez Salido, Mariano Matilla García, María Dolores Lorca y Carlos Valero, que con sus ideas y su experiencia han ayudado a que la versión de este libro haya mejorado, además de aquellos compañeros del INE que han prestado su disposición y colaboración en todo aquello que hemos necesitado.

Queremos agradecer también a Miguel Ángel de Castro y a Miguel Ángel Martínez toda su ayuda y apoyo en este proyecto, continuación del que surgió en 2021, pero que con un cambio de temario muy significativo en los procesos selectivos había que actualizar, y que ellos recibieron con entusiasmo y han impulsado para que pudiera ser una realidad.

En particular nos acordamos de nuestras familias, a las que hemos robado mucho tiempo para que esta iniciativa saliera adelante, y que siempre nos apoyan, y a tantos amigos que nos han animado con esta aventura.

Prólogo

Este libro ha nacido para que puedan utilizarlo todos aquellos opositores que estén preparando el acceso a los cuerpos de estadística del estado, aunque por su carácter multidisciplinar puede ser de utilidad en otros ámbitos.

Esperamos que tengáis la bondad de perdonar los errores que pudiéramos haber cometido al elaborarlo.

También nos gustaría agradecer las correcciones, aportaciones e ideas de quien en el uso de este libro pudiera detectar alguna cuestión que fuera susceptible de cambio con el objetivo de mejorar este material. Todo aquello que recibamos será considerado para publicar futuras versiones mejoradas y que puedan seguir siendo útiles para toda la comunidad estadística. Nuestro contacto es el siguiente:

juhersan.zz@gmail.com

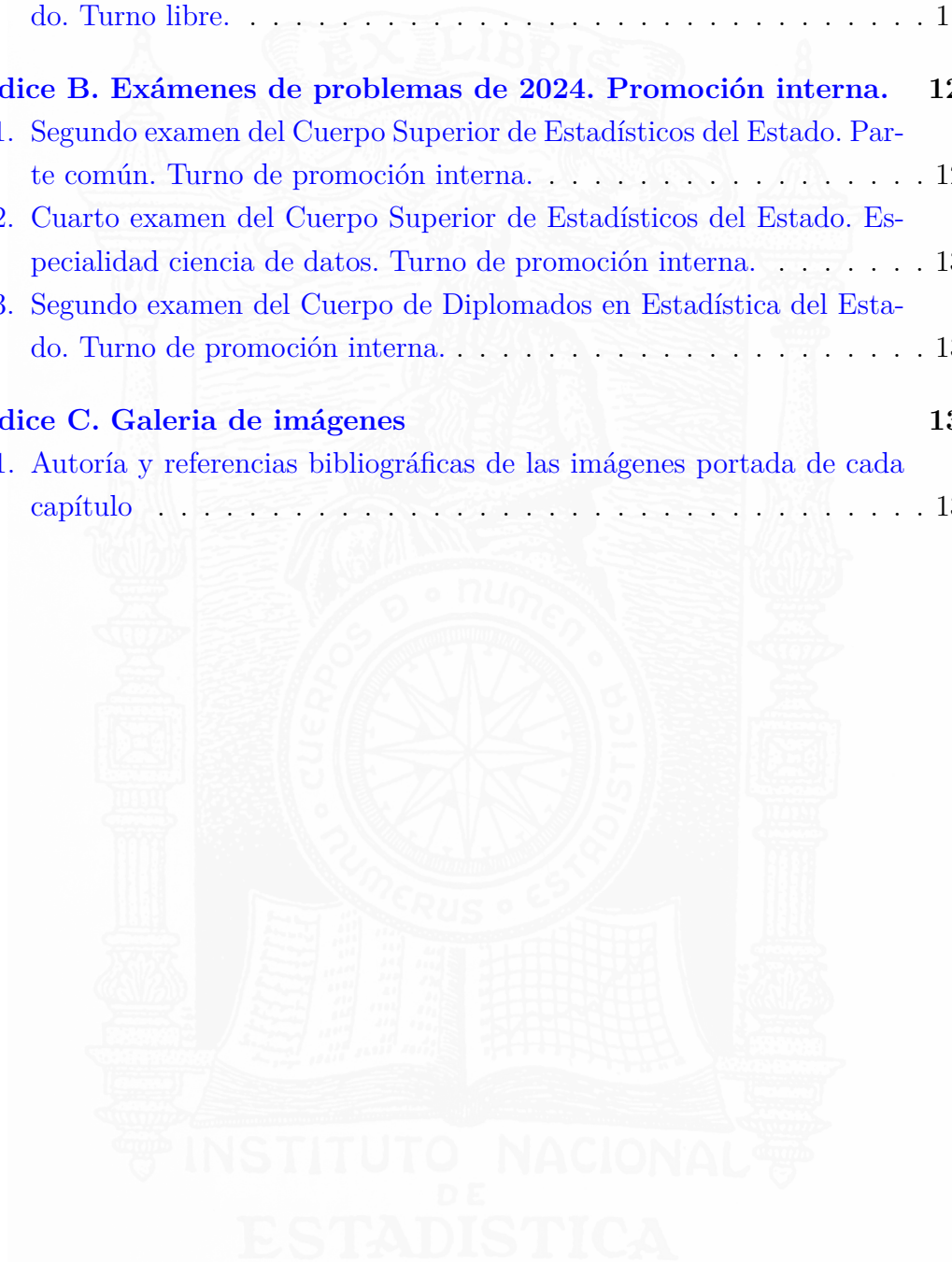
zeycus@gmail.com

Este libro se irá actualizando con los exámenes de problemas que cada año vayan publicándose de las sucesivas convocatorias del cuerpo de diplomados en estadística del estado y con los exámenes de la parte común del cuerpo superior y de la especialidad de ciencia de datos, con la idea de crear una herramienta útil y actual para el opositor/a.

Índice general

Agradecimientos	VII
Prólogo	IX
1. Introducción	1
2. Año 2024. Turno libre.	9
2.1. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del Estado de 2024.	10
2.2. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte Común.	42
2.3. Soluciones del cuarto examen del Cuerpo Superior de Estadísticos del Estado de 2024. Especialidad ciencia de datos. Turno libre.	85
3. Año 2024. Promoción interna.	87
3.1. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte común. Promoción interna.	88
3.2. Soluciones del cuarto examen del Cuerpo Superior de Estadísticos del Estado de 2024. Especialidad ciencia de datos. Promoción interna. . .	96
3.3. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del Estado de 2024. Promoción interna.	96
Bibliografía	97
APÉNDICES	
Apéndice A. Exámenes de problemas de 2024. Turno libre.	103
A.1. Segundo examen del Cuerpo Superior de Estadísticos del Estado. Parte común. Turno libre.	103

A.2. Cuarto examen del Cuerpo Superior de Estadísticos del Estado. Especialidad ciencia de datos. Turno libre.	114
A.3. Segundo examen del Cuerpo de Diplomados en Estadística del Estado. Turno libre.	114
Apéndice B. Exámenes de problemas de 2024. Promoción interna.	123
B.1. Segundo examen del Cuerpo Superior de Estadísticos del Estado. Parte común. Turno de promoción interna.	123
B.2. Cuarto examen del Cuerpo Superior de Estadísticos del Estado. Especialidad ciencia de datos. Turno de promoción interna.	132
B.3. Segundo examen del Cuerpo de Diplomados en Estadística del Estado. Turno de promoción interna.	132
Apéndice C. Galería de imágenes	139
C.1. Autoría y referencias bibliográficas de las imágenes portada de cada capítulo	139





Capítulo 1

Introducción



*“True science teaches, above all, to doubt
and to be ignorant.”*

Miguel de Unamuno



Actualmente los procesos selectivos a los Cuerpos de Estadísticos del Estado se ajustan a lo dispuesto en un Real Decreto de 2015, de 30 de octubre, por el que se aprueba el texto refundido de la Ley del Estatuto Básico del Empleado Público, y constan de una fase de oposición, tanto para los aspirantes que se presenten por el sistema general de acceso libre, como para los aspirantes que se presenten por la convocatoria de promoción interna. Además el proceso selectivo, para aquellos que superen la fase de oposición, incluye un curso selectivo que se valora y que influye en la puntuación final de los candidatos.

En la convocatoria de 2022 del cuerpo superior de estadísticos del estado el temario ha sufrido un cambio muy significativo respecto de los procesos anteriores que se venían desarrollando. A partir de este año hay dos especialidades, una relativa a Estadística-Ciencia de datos y otra relativa a Estadística aplicada a las ciencias



económicas y sociales, compartiendo ambas una parte común de temario que tendrá que ser preparada por todos los candidatos. Relativa a esa parte de materias comunes versa el primer examen, que es una prueba de contestar por escrito a veinticinco preguntas cortas, a elegir de entre treinta propuestas por el Tribunal relacionadas con los temas de dicho bloque común. Este conjunto de conocimientos comunes a ambas especialidades se concentra en temas de derecho, demografía, cuentas nacionales, inferencia y modelización, producción estadística oficial y almacenamiento y modelos de datos. El segundo examen es un examen de problemas también de este conjunto de temas comunes en el cual de 8 problemas se eligen 4 (excepto el apartado de derecho que no se incluye en este examen). El tercer examen es un examen de idiomas en el que se establecen dos modalidades excluyentes de evaluación: una presencial y, otra, sustitutiva de la anterior, mediante la cual los candidatos podrán acreditar su conocimiento de inglés equivalente a B2 o superior, de acuerdo con los niveles del Marco Común Europeo de Referencia de las Lenguas (MCER). La prueba presencial consistirá en la realización de una prueba, escrita y otra oral. Por último, el cuarto examen consistirá en contestar por escrito a cuatro cuestiones prácticas a elegir entre ocho propuestas por el Tribunal, de los temas del grupo de materias específicas según la especialidad consignada en la solicitud de cada candidato. Un aspecto muy significativo en este cambio de paradigma es que se han eliminado exámenes escritos de desarrollar diversos temas y se ha eliminado el examen oral, que tradicionalmente era el más duro con diferencia y el que requería del mayor esfuerzo para poder asimilar y exponer ante el tribunal los temas requeridos.

El contenido del temario ha supuesto un punto de inflexión en los conocimientos que se demandan que tengan los opositores, de acuerdo al conjunto de exigencias cada vez más novedosas que se van implantando en la estadística oficial y que requieren de unos perfiles diferentes a los que habitualmente estaban entrando en estos cuerpos de estadística. En la especialidad de ciencia de datos se incluyen bloques de bases de datos, programación, tecnologías web, simulación, análisis de algoritmos, entre otros que no aparecían anteriormente. Igualmente, en la especialidad más orientada a las ciencias sociales aparecen como necesarios tener conocimientos en contabilidad financiera o el sistema fiscal español, que son novedosos también. Todos los exámenes son eliminatorios de forma que sólo aquellos aspirantes que superen todos y cada uno de ellos podrán acceder a dicho cuerpo superior. Es decir, parece que esta nueva estructura confiere más importancia a la resolución de problemas que a la acumulación memorística de conceptos y exposición de los mismos de una u otra forma, por lo que todavía es más importante para los opositores poder contar con



ayudas como la que este libro intenta ofrecer.

Por el contrario, en las oposiciones al Cuerpo de Diplomados en Estadística del Estado el temario no ha sufrido grandes variaciones, con un planteamiento basado en dos exámenes, uno de preguntas cortas sobre diversos bloques de contenido y uno de problemas que engloba a todas las áreas de conocimiento del temario. La incorporación de exámenes de preguntas cortas ha sustituido a pruebas consistentes en redactar uno o varios temas de diversas áreas temáticas, pero lo que ha permanecido constante a lo largo de muchos años, y sigue vigente actualmente son los exámenes de problemas en una u otra fase de cada proceso selectivo.

Para preparar las oposiciones es necesaria la utilización de un gran número de textos de diversa índole, tanto de teoría como de problemas, pero dada la diversidad del temario no existen referencias que cubran conjuntamente aspectos de campos tan dispares como demografía, econometría, economía, estadística y muestreo. Es decir, que para profundizar en el conocimiento de cada parte se necesita bibliografía especializada. Y esta situación se traslada de igual forma a la preparación de los exámenes de problemas, teniendo que trabajar para cada apartado en la búsqueda de problemas que pudieran ser objeto selección en estas pruebas. No hay ningún material que aglutine de alguna manera el conocimiento necesario de lo que ha sido susceptible de ser preguntado y tampoco hay material que explique y resuelva los problemas en los diversos procesos de oposición a lo largo de los años. De esta forma, al opositor/a le resulta muy complicado poder tener cierta certeza de cómo se resuelven aquellas cuestiones de ramas tan diversas.

Por este motivo, este libro se enfoca en dichas pruebas, que en el acceso a ambos cuerpos corresponden a varios de los exámenes eliminatorios. En el caso del Cuerpo de Diplomados es el segundo y último examen de la oposición, en el cual se decide el proceso y cobra una importancia vital, y en el caso del Cuerpo Superior es el segundo examen de la parte común, que es la puerta al examen de idiomas y el cuarto y último examen de problemas de la parte específica.

El objetivo del libro no es otro que el de intentar que los opositores cuenten con un material de ayuda que les acompañe en todo el proceso y en su camino de preparación y que de alguna forma les genere seguridad sobre cómo abordar esta parte de la oposición, puesto que un examen de problemas no está exento de cierta incertidumbre.

Además, se han incluido en el texto los exámenes de promoción interna, tanto del cuerpo de diplomados como del cuerpo superior para poder contar con todos los exámenes disponibles de problemas, puesto que todas las modalidades de acceso





a los diferentes cuerpos aportan valor al conocimiento de la resolución de diversas cuestiones y son vitales en un momento de cambio como el que hemos comentado y del que no se disponen de problemas orientativos de diversos bloques que pueden ser susceptibles de ser preguntados en futuras convocatorias. La decisión de añadir estos exámenes se ha hecho para que el libro sea lo más compacto posible y ayude a cada aspirante desde cualquier situación administrativa que pudiera tener.

Los cuerpos de estadística del estado para los que este material pretende ser de ayuda han sufrido muchos cambios a lo largo de los últimos tres siglos. Los antecedentes se remontan a 1856 cuando se creó la Comisión Estadística General del Reino, que estaba compuesta por un grupo de especialistas en estadística que se dedicaron a supervisar, coordinar y asesorar a los diversos Departamentos de la Administración en el ámbito estadístico.

Los Cuerpos de Estadística en sus orígenes estuvieron muy vinculados a lo que hoy conocemos como Catastro, de forma que había una interrelación entre la topografía catastral, la geografía y la estadística. En 1870 se creó una Dirección General de Estadística Autónoma y se publicó un listado con los funcionarios que llevaron a cabo trabajos estadísticos y topográficos. Pero es en 1873 cuando se vio conveniente crear un Cuerpo de Estadística con diversas categorías administrativas, cómo había que acceder, así como las retribuciones y otros aspectos de relevancia.

Fue en 1902 cuando un Real Decreto del entonces Ministerio de Instrucción Pública y Bellas Artes establecía una separación del Cuerpo de Estadística en dos partes: el Cuerpo Facultativo de Estadística y el Cuerpo Auxiliar de Estadística, fijando las plantillas de ambos.

En los años treinta la estadística tuvo una gran explosión con los avances de la estadística matemática, así como de otros campos como la probabilidad y la inferencia, pero la guerra civil afectó a todo el engranaje estadístico de forma considerable. Se incorporaron, no obstante, máquinas para el procesamiento de información que hasta entonces se hacía de forma manual, contribuyendo a la formación del Censo de Población en 1930.

El Real Decreto de 29 de enero de 1930 del Ministerio de Trabajo y Previsión aprobó un reglamento que establecía las funciones de los Cuerpos Facultativo y Técnico de Ayudantes de Estadística, incluyendo aspectos sobre el ingreso, vacantes, tribunales de oposición etc...

El ingreso en el Cuerpo Facultativo de Estadística se llevaría a cabo por oposición con ciertos requisitos sobre la edad y la formación de los candidatos, teniendo que superar una serie de pruebas sobre álgebra, cálculo, estadística, economía, derecho



y geografía. Para el Cuerpo de Ayudantes de Estadística también se estableció un sistema de oposición, pero no se requería titulación académica.

En 1942 se creó el Cuerpo Técnico de Estadística y Colocación que requería para poder formar parte de él también un proceso de concurso-oposición para el que era necesario tener un título universitario. Es en la Ley Estadística de 1945 cuando queda reflejado que los trabajos del INE deberían ser llevados a cabo por dos cuerpos de nueva creación, que fueron el Cuerpo de Estadísticos Facultativos y el Cuerpo de Estadísticos Técnicos. Al primero le correspondería el trabajo de dirección y de investigación y el acceso al mismo sería mediante oposición entre titulados de grado superior. Respecto al segundo, se encargaría de la ejecución técnica de los diversos trabajos estadísticos.

En los años 60 se produjo un gran cambio en los cuerpos de estadística puesto que había que acometer nuevos retos profesionales e institucionales, a la vez que una evolución en la legislación del funcionariado también requería diversas adaptaciones. El INE inicia operaciones importantes en esos años, como la Encuesta de Población Activa(EPA), pero se producen situaciones que restan importancia al papel que la institución tenía y trataba de adquirir. Otros organismos e incluso los propios Ministerios comienzan a elaborar estadísticas de diversa índole al margen del INE, de forma que poco a poco esa labor que se supone debía tener la institución como responsable del sistema estadístico se va debilitando en cierto modo.

Para el acceso al Cuerpo de Estadísticos Facultativos, en los años 60, se requería un título universitario de grado superior, pero a partir del 68 se limitaba fundamentalmente, salvo alguna excepción en alguna convocatoria, a licenciados en matemáticas y en ciencias políticas, económicas y comerciales. El temario fue cambiando con el paso del tiempo, y se introdujeron materias como informática y econometría a las ya existentes de derecho, demografía, economía y estadística. Esta circunstancia hizo que la dificultad en las mismas fuera en aumento puesto que no se adecuaba el temario a ningún plan de las carreras universitarias existentes. La plantilla del Cuerpo era de 200 funcionarios en 1945 y se amplió en 40 plazas en 1973, con la idea a partir de 1974 de dotar de 10 por año al cuerpo, pero muchas quedaron sin cubrir por la dificultad de los conocimientos necesarios para el acceso, así como de las pruebas selectivas.

En cuanto al Cuerpo de Estadísticos Técnicos, inicialmente se requería un título de enseñanza media o asimilado, pero a partir de 1968 había que poseer un título de Bachiller Superior, Perito Mercantil o Maestro de Primaria, a la vez que se actualizaron los contenidos y exámenes de la oposición, que se centrarían en Matemáticas,





Estadística, Legislación y Economía fundamentalmente. La plantilla inicial fue de 300 funcionarios y la Ley del 73 la incrementó en 120 plazas, que se convocarían a partir de 1974 de forma gradual.

El inicio del uso de los ordenadores en el INE se remonta al año 1964, en el que se utilizó un equipo IBM con tarjetas perforadoras y 16Kb de memoria y para el que fue necesario formar a un equipo de 14 personas, entre facultativos y técnicos. En la década de los 70 se van sustituyendo esos equipos por otros más modernos y con más capacidad, que van dotando al organismo de experiencia en la evolución de las posibilidades técnicas que van ofreciendo estos avances.

En el año 77, mediante Real Decreto, se creó el Cuerpo de Estadísticos Técnicos Diplomados, extinguiéndose así el Cuerpo de Estadísticos Técnicos. Una orden en ese mismo año regularía la integración de un cuerpo en el otro.

A finales de esta década se empezó a gestar la elaboración de una nueva ley de estadística que dejase atrás la ley de 1945. Ya la Constitución del 78 establecía las competencias del Estado y entre ellas incluía la estadística para fines estatales, aunque este último término no estaba suficientemente definido. Por este motivo era necesario elaborar una nueva ley de Estadística que pudiera aclarar y desarrollar esta problemática.

Ya en el año 80 se configuraba en los distintos documentos previos a la nueva ley al INE como punto central y pilar fundamental del Sistema Estadístico Nacional, de forma que sería el Plan Nacional de Estadística el que reflejara las estadísticas con fines estatales, así como aquellas que fueran de interés para las CCAA.

En la década de los 80 se exigió ser Licenciado, Ingeniero, Arquitecto o estudios equivalentes para poder acceder al Cuerpo de Estadísticos Facultativos y al igual que ocurrió anteriormente el temario se fue adaptando progresivamente a los avances en las distintas disciplinas que lo componían, como informática y econometría.

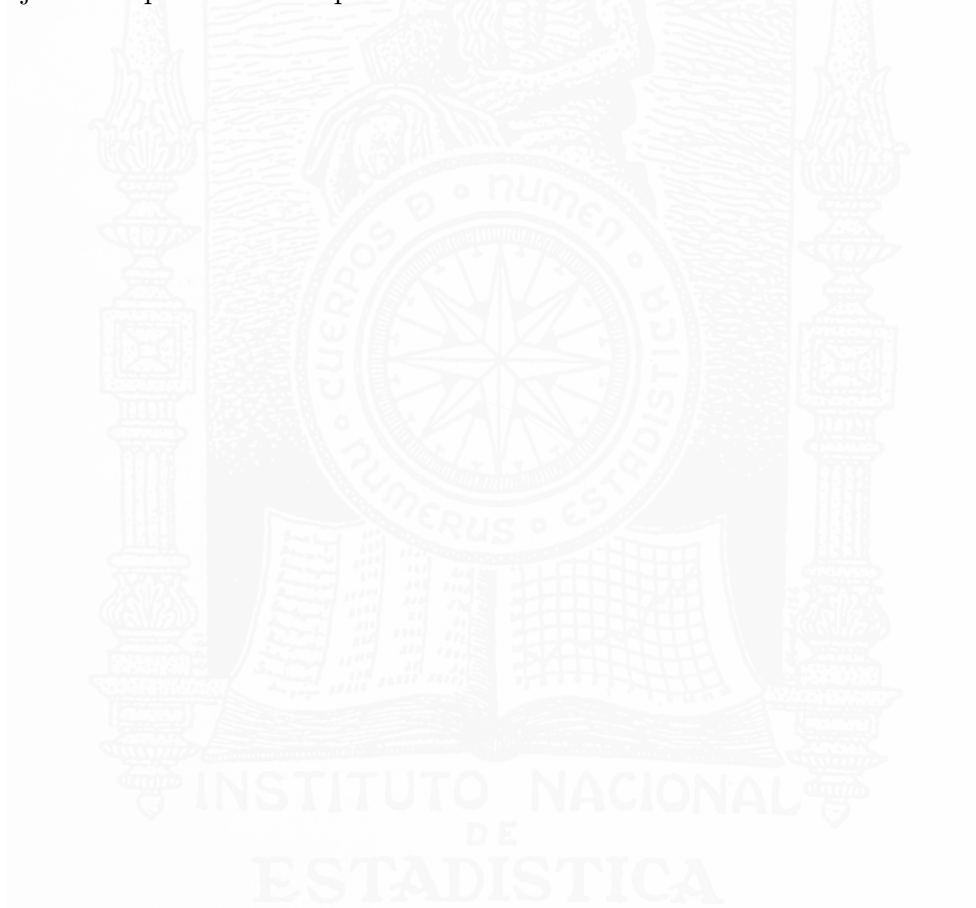
La publicación de la Ley de la Función Estadística Pública tuvo lugar en el año 89, en la cual se elevaba el rango administrativo del INE, que pasaba a ser Subsecretaría de Estado, y establecía la posibilidad de tener autonomía dentro de la Administración, de forma que pasaba a ser un Organismo Autónomo adscrito al Ministerio de Economía y Hacienda.

En el año 2001, la Ley de Medidas Fiscales Administrativas y de Orden Social cambia la denominación del Cuerpo de Estadísticos Facultativos, que a partir de ese momento se llamaría Cuerpo Superior de Estadísticos del Estado, e igualmente cambia la del Cuerpo de Estadísticos Técnicos Diplomados por el nombre Cuerpo de Diplomados en Estadística del Estado, que son los nombres actualmente vigentes de



los cuerpos sobre los que este libro de problemas hace referencia. Puede consultarse una exposición mucho más extensa y detallada de la historia de los cuerpos de estadística de la Administración General del Estado en referencias como [Celestino, 2010] y [Merediz, 2004].

Por último, hemos considerado importante añadir en los anexos de este ejemplar, para cada año, el enunciado de los exámenes de problemas (uno del cuerpo de diplomados y los dos, segundo y cuarto, del cuerpo superior, enfocado en la especialidad de ciencia de datos) tal y como se han presentado a los opositores, teniendo cada capítulo del libro el objetivo de resolver todas o la mayoría de las cuestiones propuestas con el mayor detalle posible. Además iremos actualizando el contenido de esta versión con la resolución de los exámenes que se realicen en futuras convocatorias para que este material se vaya consolidando y sea útil para tantos estudiantes cuyo objetivo es poder formar parte de estos colectivos de la administración.







Capítulo 2

Año 2024. Turno libre.

“Sometimes you have to go up really high to understand how small you are.”

—Felix Baumgartner.



Después de cinco años de planificación escrupulosa, el proyecto Red Bull Stratos de 20 millones de dólares hizo historia ese día, lo que representa un gran avance en el conocimiento mundial sobre la forma en que el cuerpo enfrenta las condiciones extremas cercanas al espacio.

A las 9:28 a.m. hora local, Felix despegó de Roswell, Nuevo México, EE.UU. Destino: el borde del espacio. En las próximas horas, Félix volvería a la Tierra convirtiéndose en el primer humano en romper la barrera del sonido en caída libre, completó el salto en paracaídas a mayor altura (38,969.4 metros) y alcanzó la velocidad más rápida en caída libre (1.357,6 km/h).

$$2024 = 2^3 \cdot 11 \cdot 23$$

2024 is a Cunningham number, because is equal to $45^2 - 1$.



2.1. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del Estado de 2024.

Cuestión práctica 1.

Durante un apagón localizado en cuatro barrios de una ciudad española se han medido dos variables: la duración del apagón (en horas) y el número de hogares afectados por barrio, tal y como se puede ver en la siguiente tabla:

Duración del apagón (en horas)	Hogares del Barrio A	Hogares del Barrio B	Hogares del Barrio C	Hogares del Barrio D
1	10	15	12	13
2	20	30	24	26
3	30	45	36	39
4	40	60	48	52
5	50	75	60	65

Teniendo en cuenta estos datos:

- 1) Calcule las frecuencias marginales de los hogares afectados por el apagón en cada barrio.
- 2) Si nos fijamos en el barrio B, ¿Cuál es la frecuencia condicionada a que el apagón dure 5 horas?
- 3) ¿Cuál es la distribución de la duración del apagón si sabemos que los hogares afectados están en el barrio A?
- 4) Calcule la duración media del apagón en horas por hogar.

Solución

1. Calculemos las frecuencias marginales de los hogares afectados por el apagón en cada barrio, que representan el total de hogares afectados en cada barrio, sumando todas las duraciones del apagón. Para obtenerlas, sumamos los hogares afectados en cada columna de la tabla:

■ **Barrio A:** $n_A = 10 + 20 + 30 + 40 + 50 = 150$ hogares



- **Barrio B:** $n_B = 15 + 30 + 45 + 60 + 75 = 225$ hogares
- **Barrio C:** $n_C = 12 + 24 + 36 + 48 + 60 = 180$ hogares
- **Barrio D:** $n_D = 13 + 26 + 39 + 52 + 65 = 195$ hogares

2. La frecuencia condicionada a que el apagón dure 5 horas en el Barrio B se refiere a la proporción de hogares afectados en el Barrio B cuando el apagón dura exactamente 5 horas. Según la tabla, en el barrio B hay 75 hogares cuya duración del apagón fue de 5 horas. La frecuencia condicionada es:

$$f(Y|_{X=\text{'Barrio B'}}) = \frac{n(\text{'Barrio B'}, Y=5)}{n(\text{'Barrio B'})} = \frac{75}{225} = \frac{1}{3}$$

3. La distribución condicional de la duración del apagón, dado que los hogares están en el Barrio A, se calcula dividiendo el número de hogares afectados en cada duración por el total de hogares en el Barrio A (150 hogares):

Duración (horas) $Y _{X=\text{'Barrio A'}}$	Hogares Barrio A $n(X = \text{'Barrio A'}, Y = y)$	Frecuencia condicional $f(Y = y _{X=\text{'Barrio A'}}) = \frac{n(X=\text{'A'}, Y=y)}{n_A}$
1	10	$\frac{10}{150} = \frac{1}{15} \approx 0,0667(6,67\%)$
2	20	$\frac{20}{150} = \frac{2}{15} \approx 0,1333(13,33\%)$
3	30	$\frac{30}{150} = \frac{1}{5} = 0,2(20\%)$
4	40	$\frac{40}{150} = \frac{4}{15} \approx 0,2667(26,67\%)$
5	50	$\frac{50}{150} = \frac{1}{3} \approx 0,3333(33,33\%)$

4. Para calcular la duración media por hogar primero calculamos el número total de hogares afectados, que son $150 + 225 + 180 + 195 = 750$ hogares.

Ahora calculamos el número total de horas en las que se prolongó el apagón:

$$1h \times (10 + 15 + 12 + 13) = 1 \times 50 = 50 \text{ horas}$$

$$2h \times (20 + 30 + 24 + 26) = 2 \times 100 = 200 \text{ horas}$$

$$3h \times (30 + 45 + 36 + 39) = 3 \times 150 = 450 \text{ horas}$$





$$4h \times (40 + 60 + 48 + 52) = 4 \times 200 = 800 \text{ horas}$$

$$5h \times (50 + 75 + 60 + 65) = 5 \times 250 = 1250 \text{ horas}$$

$$\text{Total de horas} = 50 + 200 + 450 + 800 + 1250 = 2750 \text{ horas}$$

Finalmente, la duración media se calcula como el cociente de ambos valores:

$$\text{Duración media} = \frac{\text{Total de horas}}{\text{Total de hogares afectados}} = \frac{2750}{750} = \frac{11}{3} \approx 3.6 \text{ horas}$$

Cuestión práctica 2.

El consumo de leche de una guardería y los índices de precios de la leche en un determinado país en cinco años consecutivos han sido los siguientes:

Año	Consumo en litros de leche	Índice de precios de la Leche (Base 2021 = 100)
2020	900	90
2021	1.000	100
2022	1.100	102
2023	1.200	114
2024	1.100	120

El precio de la leche en el año 2024 asciende a 2 Euros/litro.

Calcule, a precios corrientes de cada año, el gasto en leche que ha tenido dicha guardería en cada uno de los cinco años mostrados.

Solución

Lo primero que vamos a hacer es calcular el precio de la leche en cada año.

El índice de precios (base 2021 = 100) nos permite calcular los precios relativos. Sabemos que en 2024 el precio es de 2 €/litro con índice 120.

- El precio base, es decir, el precio en el año 2021 vale:

$$\text{Precio}_{2021} = \left(\frac{\text{Índice}_{2021}}{\text{Índice}_{2024}} \right) \times \text{Precio}_{2024} = \left(\frac{100}{120} \right) \times 2 = \frac{5}{3} \approx 1.6 \text{ €/litro}$$



Capítulo 2.1. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del Estado de 2024.

- Para calcular los precios para otros años podemos utilizar la siguiente relación:

$$\text{Precio}_{\text{año}} = \left(\frac{\text{Índice}_{\text{año}}}{100} \right) \times \frac{5}{3} \text{ €/litro}$$

Año	Índice de precios	Precio (€/litro)
2020	90	$0,9 \times \frac{5}{3} = 1,5$
2021	100	$\frac{5}{3} \approx 1,6$
2022	102	$1,02 \times \frac{5}{3} \approx 1,7$
2023	114	$1,14 \times \frac{5}{3} \approx 1,9$
2024	120	2,0

EL cálculo del gasto en leche a precios corrientes en cada año es sencillo, puesto que:

$$\text{Gasto del año} = \text{Consumo} \times \text{Precio del año}$$

Año	Consumo (litros)	Precio (€/litro)	Gasto total (€)
2020	900	1.5	$900 \times 1,5 = 1,350$
2021	1,000	$1,6$	$1,000 \times \frac{5}{3} \approx 1,666,67$
2022	1,100	1.7	$1,100 \times 1,7 = 1,870$
2023	1,200	1.9	$1,200 \times 1,9 = 2,280$
2024	1,100	2.0	$1,100 \times 2,0 = 2,200$





Cuestión práctica 3.

Sean X e Y dos variables aleatorias cuya función conjunta de densidad es la siguiente:

$$f_{XY}(x, y) = \begin{cases} k \cdot x, & 0 < x < y < 2 \\ 0, & \text{resto de casos} \end{cases}$$

Donde k es una constante de valor desconocido. Obtenga razonadamente:

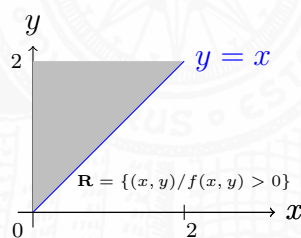
- El valor de k .
- La distribución marginal de Y .
- $P(X < 0,5 / Y = 1)$
- Una aproximación de $P(\sum_{i=1}^{100} W_i < 100)$, siendo $W_1, W_2, \dots, W_{100}$, un total de 100 variables aleatorias independientes con idéntica distribución a X .

Solución

La función de densidad conjunta tiene la expresión

$$f(x, y) = \begin{cases} k \cdot x, & 0 \leq x \leq y \leq 2, \\ 0, & \text{en otro caso} \end{cases}$$

y la región en la que vive dicha función es:



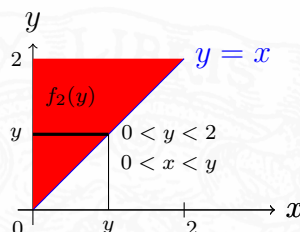
a) Tenemos que comprobar que $f(x, y)$ es una función de densidad, es decir, que integra 1 en el recinto donde vive dicha función.

$$\begin{aligned} 1 &= \int_0^2 \int_x^2 kx \, dy \, dx = \int_0^2 kx [y]_x^2 \, dx = k \int_0^2 x(2 - x) \, dx \\ &= k \int_0^2 (2x - x^2) \, dx = k \left[\frac{2x^2}{2} - \frac{x^3}{3} \right]_0^2 \\ &= k \left[4 - \frac{8}{3} \right] = k \cdot \frac{4}{3} \end{aligned}$$



Luego, $k = \frac{3}{4}$.

b) Calculamos la función de densidad marginal de Y :



$$f_Y(y) = \int_0^y \frac{3}{4} dx = \frac{3}{4} [x]_0^y = \frac{3}{4} \cdot y = \frac{3}{8} y^2, \quad 0 \leq y \leq 2.$$

c) Primeramente calculamos la función de densidad condicionada, puesto que es necesaria para obtener lo que nos piden en este apartado.

$$f(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{3}{4} \cdot x}{\frac{3}{8} \cdot y^2} = \frac{2x}{y^2}, \quad 0 \leq x \leq y$$

$$P(X \leq 0,5 | Y = 1) = \int_0^{0,5} \frac{2x}{y^2} \Big|_{y=1} dx = \int_0^{0,5} 2x dx = [x^2]_0^{0,5} = \frac{1}{4}.$$

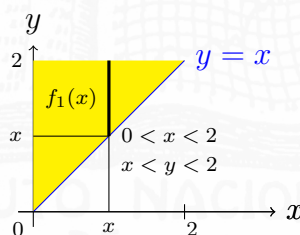
d) Sea

$$Z = \sum_{i=1}^{100} W_i, \quad \text{con } W_i \text{ variable aleatoria con igual distribución que } X.$$

Nos piden

$$P(Z \leq 100).$$

Calculamos la distribución de X :



$$f_X(x) = \int_x^2 \frac{3}{4} dy = \frac{3}{4} x [y]_x^2 = \frac{3}{4} x (2 - x), \quad 0 \leq x \leq 2.$$

Por el teorema central del límite tenemos que

$$Z = \sum_{i=1}^{100} W_i \sim N(E(Z), V(Z)) = N(nE(X), \sqrt{nV(X)}).$$





Ahora calculamos $E(X)$:

$$\begin{aligned} E(X) &= \int_0^2 x \cdot \frac{3}{4}x(2-x) dx = \frac{3}{4} \int_0^2 x^2(2-x) dx = \frac{3}{4} \left[\frac{2x^3}{3} - \frac{x^4}{4} \right]_0^2 = \frac{3}{4} \left[\frac{16}{3} - \frac{16}{4} \right] = \\ &= \frac{3}{4} \left(\frac{16}{3} - 4 \right) = \frac{3}{4} \cdot \frac{4}{3} = 1. \end{aligned}$$

El valor de la varianza de X :

$$V(X) = E(X^2) - (E(X))^2 = E(X^2) - 1.$$

Primero obtenemos $E(X^2)$:

$$\begin{aligned} E(X^2) &= \int_0^2 x^2 \cdot \frac{3}{4}x(2-x) dx = \frac{3}{4} \int_0^2 x^3(2-x) dx = \frac{3}{4} \int_0^2 (2x^3 - x^4) dx = \\ &= \frac{3}{4} \left[\frac{2x^4}{4} - \frac{x^5}{5} \right]_0^2 = \frac{3}{4} \left[\frac{32}{4} - \frac{32}{5} \right] = \frac{3}{4} \left(8 - \frac{32}{5} \right) = \frac{3}{4} \cdot \frac{8}{5} = \frac{6}{5}. \end{aligned}$$

Entonces, la varianza de X es:

$$V(X) = \frac{6}{5} - 1$$

Finalmente, aplicamos el resultado para la distribución de Z :

$$Z \sim N \left(100 \cdot 1, \sqrt{100 \cdot \frac{1}{5}} \right) = N(100, \sqrt{20}).$$

Luego,

$$\begin{aligned} P(Z \leq 100) &= P(N(100, \sqrt{20}) \leq 100) = \\ &= P\left(N(0, 1) \leq \frac{100 - 100}{\sqrt{20}}\right) = P(N(0, 1) \leq 0) = \frac{1}{2}. \end{aligned}$$



Cuestión práctica 4.

En un estudio estadístico se dispone de información de dos variables: X e Y . Se han resumido los datos en la siguiente tabla:

x_i	y_j	n_{ij}
10-30	0-10	3
10-30	10-20	3
30-50	10-20	2
50-110	20-30	2

Además, se han realizado algunos cálculos relativos a los momentos bidimensionales respecto al origen a_{rs} en relación a la citada información (r es el orden para la variable X , y s es el orden para la variable Y):

$$a_{10} = 36 \quad a_{20} = 1.840 \quad a_{01} = 14 \quad a_{02} = 245 \quad a_{11} = 640$$

Se está interesado en realizar predicciones de la variable Y utilizando la información de la variable X . Responda razonadamente a las siguientes preguntas:

a) Obtenga la recta de regresión de Y sobre X que minimice la varianza residual.

b) Considere las variables Z y W , que se definen como sigue:

$$Z=2X$$

$$W=Y+3$$

Calcule la recta de regresión de W sobre Z que minimice la varianza residual.





Solución

(a) La recta de regresión de Y sobre X es:

$$y - \bar{y} = \frac{S_{xy}}{S_x^2} (x - \bar{x})$$

$$\bar{y} = a_{10} = 14$$

$$\bar{x} = a_{01} = 36$$

$$S_{xy} = a_{11} - a_{10}a_{01} = 640 - 14 \cdot 36 = 136$$

$$S_x^2 = a_{20} - a_{10}^2 = 1840 - 36^2 = 544$$

Luego,

$$y - 14 = \frac{136}{544} (x - 36)$$

y despejamos:

$$y = \frac{136}{544}x - \frac{36 \cdot 136}{544} + 14 = \frac{136}{544}x - 9 + 14 = \frac{x}{4} + 5$$

(b)

$$Z = 2X$$

$$W = Y + 3$$

$$E(Z) = 2E(X) = 2 \cdot 36 = 72$$

$$V(Z) = 4V(X) = 4 \cdot 544 = 2176$$

$$E(W) = 3 + E(Y) = 3 + 14 = 17$$

$$V(W) = V(Y) = E(Y^2) - (E(Y))^2 = a_{02} - (a_{01})^2 = 245 - 14^2 = 49$$

$$S_{ZW} = \text{Cov}(Z, W) = \text{Cov}(2X, Y + 3) = 2 \cdot \text{Cov}(X, Y) + \text{Cov}(2X, 3)$$

$$S_{ZW} = 2S_{XY} + 0 = 2 \cdot 136 = 272$$

La recta de W sobre Z que minimiza la varianza residual se obtiene utilizando el método de mínimos cuadrados y tiene por expresión:

$$w - \bar{w} = \frac{S_{WZ}}{S_Z^2} (z - \bar{z})$$

$$w - 17 = \frac{S_{ZW}}{S_W^2} (z - 72)$$

$$w - 17 = \frac{272}{2176} (z - 72)$$

Despejamos:

$$w = \frac{Z}{8} - \frac{272 \cdot 72}{2176} + 17 = \frac{z}{8} + 8$$



Capítulo 2.1. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del
Estado de 2024.





Cuestión práctica 5.

Un grupo de científicos militares trabajan desarrollando una bomba antes de que lo hagan sus enemigos. Los datos recolectados en 25 pruebas experimentales sugieren que la energía liberada (en kilotones) sigue una distribución normal de media y varianza desconocida.

Las mediciones obtenidas para la media y cuasidesviación típica muestral son las siguientes:

$$\bar{x} = 20,75 \text{ kilotones}$$

$$s_c = 1,25 \text{ kilotones}$$

- Calcule un intervalo de confianza del **95%** para la media de la energía liberada en las explosiones experimentales.
- Se necesita que la energía media liberada supere los **20 kilotones**. Realice un contraste de hipótesis con un nivel de significación del **5%**. ¿Deben los científicos continuar perfeccionando de la bomba, o la energía liberada es suficiente?
- Calcule la región de rechazo para la media, $\bar{x}$, con un nivel de significación de $\alpha = 0,01$. Tenga en cuenta que ahora se sabe que la desviación típica poblacional es $\sigma = 1,5$. La región de rechazo tiene la forma:
$$RC = \{Z > Z_\alpha\}$$
- Si exigimos que el valor superior de la función de potencia sea de **0,0013**, ¿qué media $\bar{x}$ de energía liberada necesitamos alcanzar en **4 pruebas extra** que se vuelven a realizar?

NOTA: Se proporcionan los siguientes valores que pueden ser de utilizad para la resolución de esta cuestión práctica.

Valores tabla t de Student

$$t_{24;0,025} = 2,064; t_{24;0,05} = 1,711; t_{25;0,025} = 2,060; t_{25;0,05} = 1,708$$

Valores tabla Normal

$$Z_{0,05} = 1,64; Z_{0,025} = 1,96; Z_{0,01} = 2,33; Z_{0,0013} = 3$$



Solución

- a) El intervalo de confianza del 95 % para la media, cuando la varianza es desconocida, viene determinado por la expresión.

$$IC = \bar{x} \pm t_{n-1, \alpha/2} \cdot \frac{s_c}{\sqrt{n}}$$

siendo cada elemento de la fórmula:

- Media muestral ($\bar{x}$) = 20.75 kilotones
- Cuasidesviación típica (s_c) = 1.25 kilotones
- Tamaño muestral (n) = 25 pruebas
- Nivel de confianza = 95 % $\rightarrow \alpha = 0,05$

El Valor crítico $t_{24, 0,025} = 2,064$ se obtiene de la tabla de la distribución t-Student, siendo el valor del error estándar $= \frac{1,25}{\sqrt{25}} = 0,25$.

Por lo tanto el intervalo es:

$$20,75 \pm 0,516 \Rightarrow (20,234, 21,266) \text{ kilotones}$$

- b) El contraste de hipótesis, que es unilateral, se escribe:

- $H_0 : \mu \leq 20$ (la energía no supera 20 kilotones)
- $H_1 : \mu > 20$ (la energía supera 20 kilotones)

La región crítica en este tipo de contrastes es de la forma $RC = \{\bar{x} > c\}$. Llamando T al estadístico de contraste:

$$T = \frac{\bar{x} - \mu_0}{s_c / \sqrt{n}}$$

que sigue una distribución t de Student con $n - 1$ grados de libertad (puesto que no se conoce la varianza de la población) dicha región crítica se convierte en:

$$RC = \{\bar{x} > c\} = \left\{ T > \frac{c - \mu_0}{s_c / \sqrt{n}} = t_{24, 0,05} = 1,711 \right\}$$

teniendo en cuenta que utilizando un 5 % de nivel de significación sabemos que $t_{24, 0,05} = 1,711$, y por lo tanto, simplificando, la región crítica se expresa como:

- Rechazar H_0 si $T > 1,711$.





Como el valor del estadístico de contraste es $T = \frac{20,75-20}{1,25/5} = 3 > 1,711$, entonces se rechaza H_0 , por lo que parece que la energía media liberada, con los datos que tenemos, no es menor de 20 kilotones y por lo tanto no es necesario seguir perfeccionando la bomba.

c) En este caso, las hipótesis de trabajo son:

- $H_0 : \mu \leq \mu_0$
- $H_1 : \mu > \mu_0$

La región crítica en este caso es $RC = \{\bar{x} > k\} \equiv \{Z > z_\alpha\}$, con Z el estadístico de contraste, que teniendo en cuenta que ahora tenemos una población normal con σ conocida, sigue una distribución normal de media cero y varianza 1, y vale:

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

siendo la región crítica, o región de rechazo, como indica el enunciado:

$$RC = \{Z > z_{0,01}\} = \{Z > 2,33\} = \{\bar{x} > \mu_0 + 2,33 \frac{\sigma}{\sqrt{n}}\}$$

d) Vamos a calcular la función de potencia para analizar su expresión. Por definición, la función de potencia es la probabilidad de rechazar la hipótesis nula cuando el verdadero valor del parámetro es μ , es decir:

$$\beta(\mu) = P_\mu(X \in RC) = P_\mu(\bar{x} > c)$$

En el caso de que trabajemos con un nivel de significación del 5 % el valor de esa constante se calcula como:

$$\frac{c - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{c - 20}{\frac{1,25}{5}} = 1,64 = z_{0,05}$$

es decir, que $c = 20,41$, luego la función de potencia es:

$$\beta(\mu) = P_\mu(X \in RC) = P_\mu\left(N(0,1) > \frac{20,41 - \mu}{\frac{1,25}{\sqrt{n}}}\right)$$

Si representamos esa función de potencia para $n = 25$ tenemos la siguiente expresión:

$$\beta(\mu) = P_\mu\left(N(0,1) > \frac{20,41 - \mu}{\frac{1,25}{\sqrt{25}}}\right) = P_\mu(N(0,1) > 81,64 - 4\mu)$$

Si $\mu = 20$ tenemos que la función de potencia lógicamente vale el nivel de significación $\alpha = 0,05$ puesto que:

$$\begin{aligned}\beta(20) &= P_{\mu}(N(0, 1) > 81,64 - 4 \cdot 20) = P_{\mu}(N(0, 1) > 81,64 - 80) = \\ &= P_{\mu}(N(0, 1) > 1,64) = 0,05 = \alpha\end{aligned}$$

y esta función de potencia es creciente en μ claramente, puesto que cuanto mayor es esa media poblacional más negativo es el término $81,64 - 4\mu$ y por tanto esa probabilidad crece, como lo vemos en la figura 2.1.

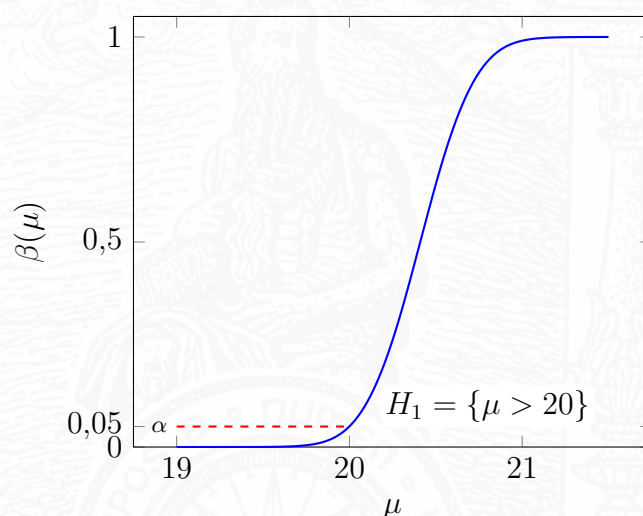


Figura 2.1: Función de potencia: $\beta(\mu) = P(N(0, 1) > 81,64 - 4\mu)$

Como puede verse, la función de potencia es siempre creciente y para $\mu = 20$ toma el valor del nivel de significación.

El enunciado habla de exigir que el valor superior de la función de potencia sea 0,0013, entendido en el sentido en el que se interpreta el valor de α , como una cota superior de la función de potencia en la hipótesis nula. De hecho, el tamaño de un contraste, por definición es el supremo de la función de potencia en la hipótesis nula, es decir, que en este caso la interpretación de este apartado es la siguiente:

$$\text{tamaño(test)} = \sup_{\mu \in H_0} \beta(\mu) = 0,0013$$

Por tanto, usando una muestra de tamaño 4, que son el número de pruebas extra, tendríamos:

$$\sup_{\mu \in H_0} \beta(\mu) = \sup_{\mu \in H_0} P_{\mu \in H_0}(\bar{x} > c) = P\left(\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > \frac{c - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) = 0,0013$$



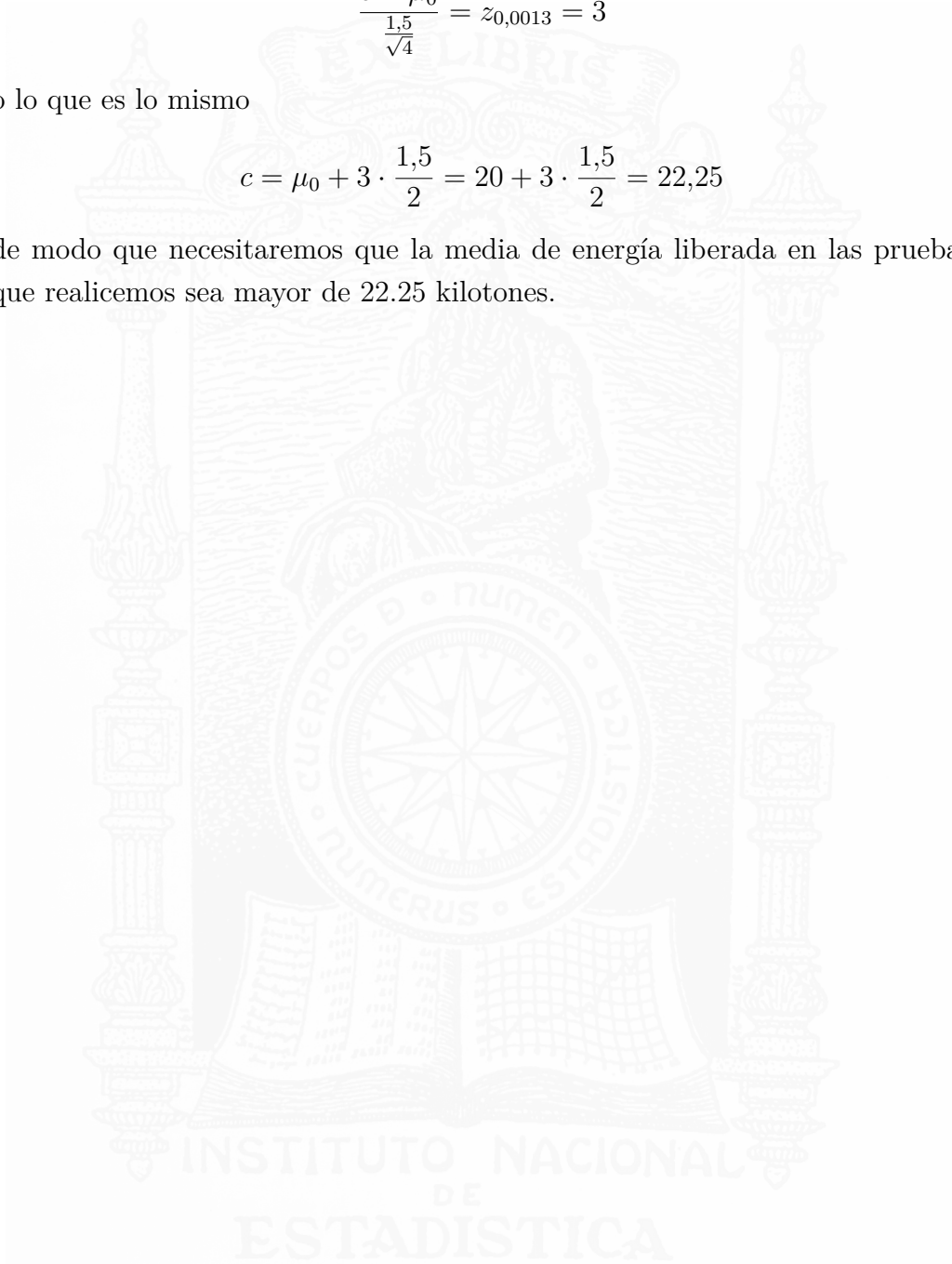
puesto que en la hipótesis nula el valor de la función de potencia es creciente y por tanto el supremo se alcanza en el valor de $\mu = \mu_0 = 20$, es decir, que

$$\frac{c - \mu_0}{\frac{1,5}{\sqrt{4}}} = z_{0,0013} = 3$$

o lo que es lo mismo

$$c = \mu_0 + 3 \cdot \frac{1,5}{2} = 20 + 3 \cdot \frac{1,5}{2} = 22,25$$

de modo que necesitaremos que la media de energía liberada en las pruebas que realicemos sea mayor de 22.25 kilotones.





Cuestión práctica 6.

Se dispone de la siguiente información de una economía (en unidades monetarias):

	TOTAL u. m.
Producción	1.700
Consumos intermedios	700
Remuneración de los asalariados	525
Remuneración de los asalariados residentes que trabajan para unidades no residentes	17
Remuneración de los asalariados no residentes que trabajan para unidades residentes	7
Excedente Bruto de Explotación / Renta mixta	425
Impuestos	
- Impuestos sobre los productos	400
- Impuestos sobre la producción y las importaciones	350
Subvenciones	
- Otras subvenciones a los productos	300
- Subvenciones a la producción y las importaciones	200
Rentas netas de la propiedad pagadas al resto del mundo	15
Consumo final efectivo individual	450
Consumo final efectivo colectivo	350

Se pide calcular:

- Valor Añadido Bruto a precios básicos.
- Producto Interior Bruto a precios de mercado de la economía. Debe obtenerse utilizando el Valor Añadido Bruto del apartado anterior.
- Renta Nacional Bruta.
- Gasto en Consumo Final.
- Completar la Cuenta de producción y la Cuenta de explotación de la economía, incluyendo sus saldos.



Solución

- a) El valor añadido bruto(VAB) a precios básicos se calcula como la diferencia entre la producción y los consumos intermedios, es decir:

$$VAB_{pb} = \text{Producción} - \text{Consumos Intermedios} = 1700 - 700 = 1000$$

- b) El producto interior bruto(PIB) a precios de mercado para el total de la economía se puede calcular utilizando el valor añadido bruto a precios básicos mediante la siguiente relación:

$$\begin{aligned} PIB_{pm} &= VAB_{pb} + \text{Impuestos productos} - \text{Subvenciones productos} = \\ &= 1000 + 400 - 300 = 1100 \end{aligned}$$

- c) Para calcular la renta nacional bruta a partir del agregado anterior es necesario corregirlo mediante las rentas primarias netas, que están compuestas por las rentas del trabajo y las rentas de la propiedad.

$$\begin{aligned} \text{Rentas trabajo netas} &= \text{R.A. residentes recibidas del exterior} - \text{R.A. no residentes} \\ &\text{pagadas al exterior} = 17 - 7 = 10 \end{aligned}$$

Respecto a las rentas de la propiedad, el problema nos da el dato de las rentas netas de la propiedad pagadas al resto del mundo, que es 15, por lo que entendemos que las rentas netas, que son las recibidas menos las pagadas tendría el valor de -15.

$$\begin{aligned} \text{Rentas primarias netas} &= \text{Rentas trabajo netas} + \text{Rentas de la propiedad netas} = \\ &= 10 - 15 = -5 \end{aligned}$$

$$RNB = PIB_{pm} + \text{Rentas primarias netas} = 1100 - 5 = 1095$$

- d) El gasto en consumo final(GCF) se compone del gasto en consumo individual y gasto en consumo colectivo. El gasto en consumo colectivo coincide con el gasto en consumo efectivo colectivo, que es un gasto que sólo tiene el sector administraciones públicas. Por otro lado, el gasto en consumo individual es un gasto no sólo del sector hogares, sino que los sectores administraciones públicas e instituciones sin fines de lucro también tienen gasto en consumo individual, que se transfiere al sector hogares en forma de transferencias sociales en especie. La suma de estos tres componentes de gasto en consumo individual conforman



el gasto en consumo individual efectivo. Ambos datos nos los proporciona el problema, por lo que tenemos:

$$\begin{aligned} GCF &= \text{Consumo efectivo individual} + \text{Consumo efectivo colectivo} = \\ &= 450 + 350 = 800 \end{aligned}$$

- e) En la figura 2.2 se muestran las cuentas de producción y de explotación detalladas.

nota: Hemos resuelto el problema con los datos que proporciona, pero no es posible que los impuestos sobre los productos sean menores que los impuestos sobre la producción y las importaciones, porque aquellos forman parte de estos. En terminos netos sí parece que es posible que sean válidos los datos, pero probablemente ha habido un error en la presentación de los datos y el valor 400 seguramente fuera el D.2 y el 350 el D.21. Con las subvenciones ocurre igual y da la sensación que están intercambiados los dos datos.



Cuentas corrientes			
I. Cuenta de producción			
Empleos	Código	Operaciones y otros flujos y saldos contables	Recursos
S.1			S.1
Total de la economía			Total de la economía
	P.1	Producción	1.700
700	P.2	Consumos intermedios	
	D.21	Impuestos sobre los productos	400
	D.31	Subvenciones sobre los productos	300
1.100	B.1*b	Producto interior bruto	
II. 1.1 Cuenta de explotación			
Empleos	Código	Operaciones y otros flujos y saldos contables	Recursos
S.1			S.1
Total de la economía			Total de la economía
	B.1*b	Producto interior bruto	1.100
525	D.1	Remuneración de los asalariados(Interior: 525)	
350	D.2	Impuestos sobre la producción y las importaciones	
200	D.3	Subvenciones sobre la producción y las importaciones	
425	B.2b/B.3b	Excedente de explotación bruto/Renta mixta bruta	
II. 1.2 Cuenta de asignación de la renta primaria			
Empleos	Código	Operaciones y otros flujos y saldos contables	Recursos
S.1			S.1
Total de la economía			Total de la economía
	B.2b/B.3b	Excedente de explotación bruto/Renta mixta bruta	425
	D.1	Remuneración de los asalariados(nacional = 525-7+17 = 535)	535
	D.2 - D3	Impuestos menos subvenciones sobre la producción y las importaciones	150
	D.4	Rentas de la propiedad netas procedentes de RM	-15
1.095	B.5*b	Renta nacional bruta	

Figura 2.2: Cuentas de producción, explotación y asignación de la renta primaria del total de la economía



Cuestión práctica 7.

Considere un monopolista con costes totales $C(q) = 2q + 1$ que se enfrenta a una curva de demanda $P(q) = 10 - q$.

- Halle el equilibrio (precio y cantidad) del monopolio.
- Halle el excedente total del monopolio y la pérdida irrecuperable de eficiencia.
- Halle la función de costes medios e indique los valores de q para que este monopolio sea considerado un monopolio natural.
- La autoridad gubernamental propone una regla de fijación de precios consistente en igualar el precio al coste marginal. Cuantifique cómo varía el excedente total del mercado como consecuencia de dicha política.

Solución

a) En un monopolio, el equilibrio se alcanza cuando el ingreso marginal (IMg) es igual al costo marginal (CMg).

Dada la función de costos totales $C(q) = 2q + 1$, el costo marginal es la derivada del costo total respecto a q :

$$CMg(q) = \frac{dC(q)}{dq} = 2$$

La función de demanda inversa es $P(q) = 10 - q$. El ingreso total (IT) es:

$$IT(q) = P(q) \cdot q = (10 - q) \cdot q = 10q - q^2$$

El ingreso marginal es la derivada del ingreso total respecto a q :

$$IMg(q) = \frac{dIT(q)}{dq} = 10 - 2q$$

Escribimos la condición de equilibrio igualando el ingreso marginal al costo marginal:

$$IMg = CMg$$

$$10 - 2q = 2 \implies 2q = 8 \implies q = 4$$

Sustituimos la cantidad $q = 4$ en la función de demanda:

$$P(4) = 10 - 4 = 6$$

obteniendo así el precio de monopolio, siendo por tanto el punto de equilibrio, como puede verse en la figura 2.3, el punto $(q_e, p_e) = (4, 6)$.



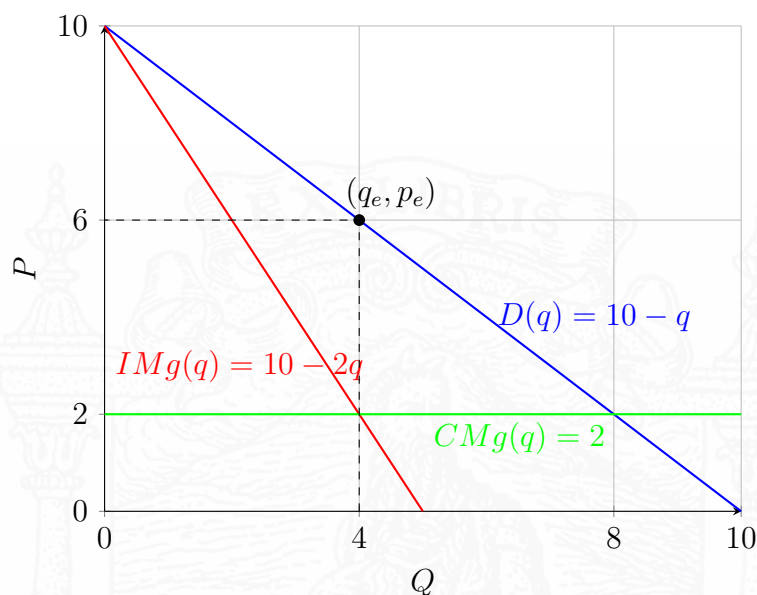


Figura 2.3: Equilibrio del monopolio.

b) El excedente total se calcula como la suma del excedente del consumidor (EC) y el excedente del productor (EP), y la pérdida irrecuperable de eficiencia (PIE) se define como la pérdida de bienestar social debido a la producción por debajo del nivel competitivo.

1. El Excedente del Consumidor (EC) es área bajo la curva de demanda menos el área determinada bajo el valor del precio de monopolio:

$$EC = \int_0^4 (10 - q) dq - 6 \cdot 4 = \left[10q - \frac{q^2}{2} \right]_0^4 - 24 = (40 - 8) - 24 = 8$$

2. El Excedente del Productor (EP) se representa como el área entre el precio de monopolio y el costo marginal:

$$EP = (6 - 2) \cdot 4 = 16$$

3. El Excedente Total (ET) es la suma de ambos, es decir:

$$ET = EC + EP = 8 + 16 = 24 \stackrel{not}{=} ET^m$$

4. La pérdida irrecuperable de eficiencia (PIE) se obtiene como la diferencia entre el excedente total en competencia perfecta y el excedente total en monopolio.



- El equilibrio competitivo equivale a: $P = CMg \Rightarrow 10 - q = 2 \Rightarrow q^c = 8, p^c = 2$.

- El excedente total competitivo se obtiene:

$$ET^c = \int_0^8 (10 - q) dq - 2 \cdot 8 = \left[10q - \frac{q^2}{2} \right]_0^8 - 16 = (80 - 32) - 16 = 32$$

- La PIE, por tanto, es:

$$PIE = ET^c - ET^m = 32 - 24 = 8$$

En la figura 2.4 tratamos de ilustrar la situación desarrollada en este apartado.

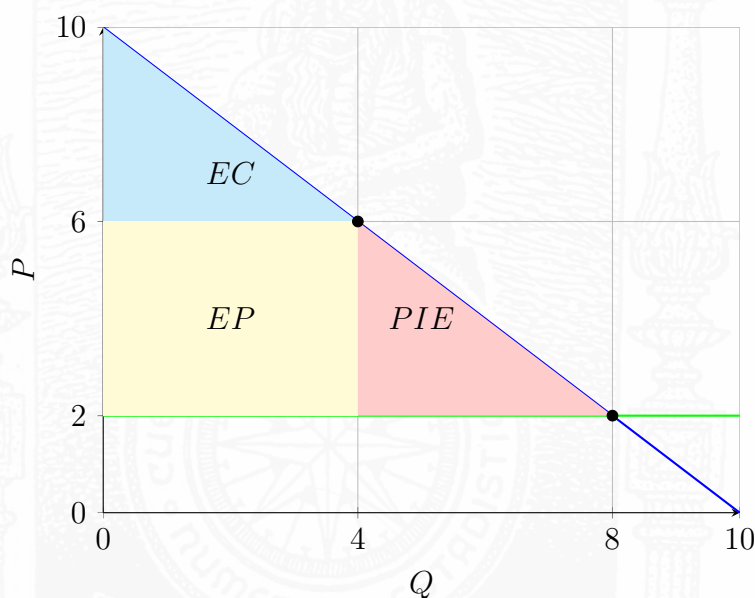


Figura 2.4: Excedente total del monopolio y pérdida irrecuperable de eficiencia.

- c) La función de costes medios(CMe) es:

$$CMe(q) = \frac{C(q)}{q} = \frac{2q + 1}{q} = 2 + \frac{1}{q}$$

Para que sea un monopolio natural, los costes medios deben ser decrecientes en q . Calculamos la derivada:

$$\frac{dCMe(q)}{dq} = -\frac{1}{q^2} < 0 \quad \forall q > 0$$

Por lo tanto, el monopolio es natural para todo $q > 0$, como puede verse en la siguiente figura 2.5.



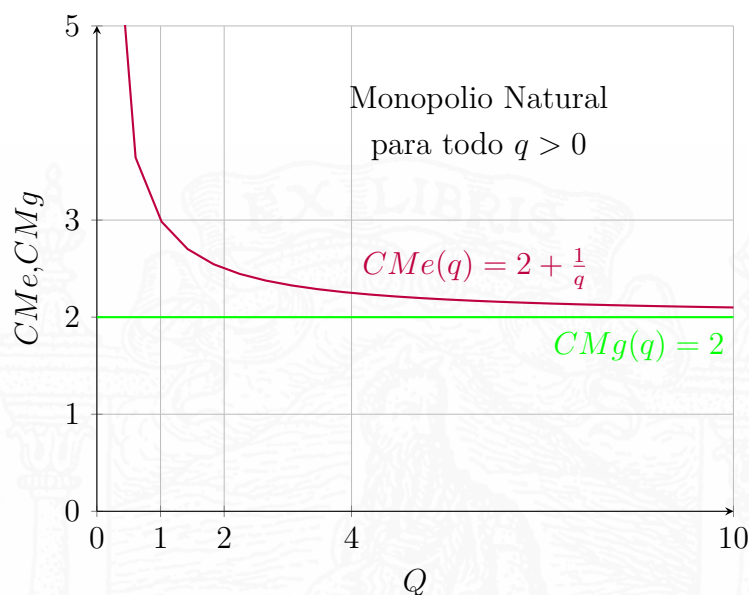


Figura 2.5: Representación del monopolio natural.

d) El enunciado postula que la regla de fijación de precios consiste en que se iguale el precio al coste marginal, es decir, $P = CM_g$.

Esta condición equivale a que el nuevo equilibrio es: $10 - q = 2 \implies q = 8$, $p = 2$.

Para calcular el excedente total (ET) procedemos como antes, calculando el excedente del consumidor y del productor y sumándolos:

$$EC = \int_0^8 (10 - q) dq - 2 \cdot 8 = 80 - 32 - 16 = 32$$

Como $P = CM_g$, el EP es cero, puesto que el precio cubre exactamente los costos marginales, pero no los fijos. Por tanto, el excedente total vale 32.

Sin embargo, en este caso, los costos totales son $C'(8) = 17$, mientras que el ingreso total es 16. Hay una pérdida de 1 para el productor, pero el excedente total sigue siendo 32 (considerando que el gobierno podría subsidiar la pérdida).

El excedente total aumenta de 24 a 32, es decir, un aumento de 8, como puede apreciarse en la figura 2.6.

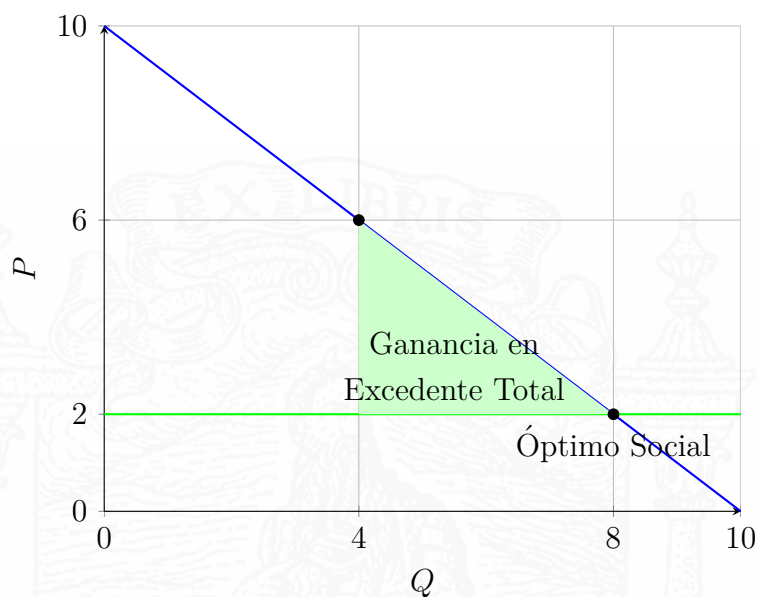


Figura 2.6: Ganancia en excedente total con la intervención de la autoridad gubernamental.



Cuestión práctica 8.

En la oficina estadística de un país europeo se dispone de los siguientes datos:

	HOMBRES					MUJERES				
	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años
Población media residente en 2024 en miles de personas	9.500	1.900	4.750	1.900	950	10.000	1.800	4.300	2.300	1.600
Defunciones de residentes registradas durante el año 2024	78.755	380	2.375	9.500	66.500	88.370	180	1.290	6.900	80.000

Teniendo en cuenta estos datos, se pide:

1. Calcule las tasas específicas de mortalidad por grupos de edad y sexo.
2. Usando las tasas del apartado anterior, calcule la tasa bruta de mortalidad por sexo.
3. Si tenemos en cuenta las tasas específicas de mortalidad del primer apartado y que la tasa bruta de mortalidad en mujeres es superior a la tasa bruta de mortalidad en hombres, explique cómo se comporta la mortalidad entre ambos sexos. ¿Por qué la mortalidad diferencial por sexo no tiene la misma tendencia en las tasas brutas por sexo que en las tasas específicas por sexo y edad? Explique una mejor forma de hacer la comparativa de las tasas por sexo en este caso.

NOTA: Puede dejar los cálculos de las tasas sin llegar a sus cifras decimales siempre que se indique de forma clara y concisa y cómo ha llegado hasta ellos.



Solución

1. Las tasas específicas de mortalidad se calculan como:

$$\text{Tasa específica}_x = \left(\frac{\text{Número de defunciones}_x}{\text{Población media residente}_x} \right) \times 1000$$

En la siguiente tabla pueden verse calculadas las tasas, para lo cuál hay que tener en cuenta que la población está en miles y que las tasas se expresan en tanto por mil.

Sexo	Grupo de edad	Población (miles)	Defunciones	TEM _x (‰)	Cálculo
Hombres	≤ 14	1900	380	0,20	(380/1900)
	15 – 44	4750	2375	0,50	(2375/4750)
	45 – 64	1900	9500	5,00	(9500/1900)
	≥ 65	950	66500	70,00	(66500/950)
Mujeres	≤ 14	1800	180	0,10	(180/1800)
	15 – 44	4300	1290	0,30	(1290/4300)
	45 – 64	2300	6900	3,00	(6900/2300)
	≥ 65	1600	80000	50,00	(80000/1600)

2. La tasa bruta de mortalidad(TBM) puede obtenerse como el promedio ponderado de las tasas específicas(expresadas en tanto por uno), utilizando el dato de la población en cada grupo de edad para establecer los pesos, y multiplicado por 1000, para expresarla en tanto por mil, es decir:

$$\text{TBM} = \frac{\sum_x (\text{Tasa específica}_x \times \text{Población}_x)}{\text{Población total}} \times 1000$$

Si las tasas se utilizan en tanto por mil no habría que multiplicar por 1000 al estar ya expresado como queremos.

Para hombres:

$$\begin{aligned} \text{TBM} &= \frac{(0,20 \times 1900) + (0,50 \times 4750) + (5,00 \times 1900) + (70,00 \times 950)}{9500} \\ &= \frac{380 + 2375 + 9500 + 66500}{9500} = \frac{78755}{9500} = 8,29\text{‰} \end{aligned}$$

Para mujeres:

$$\begin{aligned} \text{Tasa bruta} &= \frac{(0,10 \times 1800) + (0,30 \times 4300) + (3,00 \times 2300) + (50,00 \times 1600)}{10000} \\ &= \frac{180 + 1290 + 6900 + 80000}{10000} = \frac{88370}{10000} = 8,84\text{‰} \end{aligned}$$





3. Respecto al análisis de la mortalidad diferencial por sexo, la paradoja observada o la aparente contradicción es que aunque los hombres tienen mayor mortalidad específica en casi todos los grupos, la tasa bruta es más alta en mujeres. Esto se debe a que en la estructura por edad de la población femenina hay más mujeres en edades avanzadas (65+ años), donde la mortalidad es muy elevada (50‰), lo que arrastra la tasa bruta hacia arriba.

¿Por qué la tendencia no coincide? Podemos decir que hay un efecto de la composición demográfica, puesto que las mujeres tienen una pirámide poblacional más envejecida: el 16 % de ellas tiene 65+ años (vs. 10 % en hombres). Aunque su mortalidad específica en ancianas es menor, el peso proporcional de este grupo en la población femenina es mayor, aumentando la tasa bruta.

Respecto a la interpretación de las tasas brutas respecto de las tasas específicas podemos decir que las tasas brutas no ajustan por edad, por lo que son sensibles a la distribución poblacional, mientras que las tasas específicas permiten comparaciones, podríamos decir, limpias, por grupo de edad, pero ocultan el impacto global de la estructura demográfica.

El mejor método de comparación se basa en la utilización de tasas estandarizadas para evitar distorsiones por diferencias en la estructura de edad, debiéndose usar la tasa de mortalidad estandarizada por edad, para lo cuál hay que tomar una población de referencia (por ejemplo, la población total combinada) y aplicar las tasas específicas de cada sexo a esta población estándar. De esta forma se calcularían las defunciones esperadas y se podría obtener la tasa ajustada.

Cuestión práctica 9.

Consideramos los sistemas de representación de números enteros magnitud y signo, complemento a uno y complemento a dos, con 8 bits.

- Explique las diferencias entre estos tres métodos.
- Represente el número entero **-93** con los tres métodos mencionados.
- En un sistema binario (sin bit para signo), ¿A qué valor decimal exacto equivale el número **101010,101**?



Solución

a) Vamos a enumerar las principales características de los diferentes métodos de representación

1. Magnitud y signo:

- El bit más significativo (MSB) representa el signo (0 = positivo, 1 = negativo)
- Los 7 bits restantes representan la magnitud
- Rango: -127 a +127
- Tiene dos representaciones para el cero: +0 (00000000) y -0 (10000000)

2. Complemento a uno:

- El MSB sigue representando el signo
- Para números negativos, se invierten todos los bits de la versión positiva
- Rango: -127 a +127
- También tiene dos representaciones para el cero: 00000000 (+0) y 11111111 (-0)

3. Complemento a dos:

- Método más utilizado actualmente
- Para números negativos: complemento a uno + 1
- Rango ampliado: -128 a +127
- Solo una representación del cero (00000000)
- Elimina la ambigüedad del cero negativo

b) Representamos el número -93 utilizando 8 bits con los tres procedimientos detallados anteriormente:

1. Magnitud y signo:

- Valor absoluto: $93 \rightarrow 1011101$ (7 bits)
- Representación: 1 (signo negativo) + 1011101 = 11011101

2. Complemento a uno:





- Versión positiva: 01011101
- Invertir bits: 10100010

3. Complemento a dos:

- Tomar complemento a uno del valor absoluto del número, que es el 93(01011101): 10100010
- Sumar 1: $10100010 + 1 = 10100011$

c) Vamos a realizar la conversión del número binario 101010,101₂ a decimal. Ese número binario se representa mediante un sistema posicional con parte entera y fraccionaria, que vamos a detallar a continuación:

1. Parte entera (101010):

$$1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 32 + 0 + 8 + 0 + 2 + 0 = 42$$

2. Parte fraccionaria (,101):

$$1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 0,5 + 0 + 0,125 = 0,625$$

3. Valor del número:

$$101010,101_2 = 42,625_{10}$$



Cuestión práctica 10.

El propietario de una finca de naranjos está preocupado por los daños causados por una DANA, y para evaluar dicho desastre decide contratar los servicios de dos estadísticos, A y B, que realizarán un muestreo.

El dueño de la finca les proporciona la siguiente tabla, estratificada según la variedad:

Estratos (variedad)	Número de árboles de la población (N_h)	Proporción estimada de árboles afectados
1	10.000	0,7
2	6.000	0,4
3	4.000	0,15

Si el estadístico A es partidario del muestreo estratificado sin reemplazamiento con afijación proporcional, calcule:

- El tamaño de muestra total que debería utilizar para tener la misma precisión $e = 0,01$ con la que estimaría P en una muestra aleatoria simple sin estratificar.
- El tamaño muestral de cada estrato.

Por otra parte, el estadístico B prefiere el muestreo estratificado con reemplazamiento y afijación óptima. Sin embargo, su material de trabajo lo alquila.

Calcula que, para visitar cada estrato, los costes serán:

$$C_1 = 4 \text{ € por árbol}$$

$$C_2 = 4 \text{ € por árbol}$$

$$C_3 = 9 \text{ € por árbol}$$

- Calcular el tamaño muestral de cada estrato, si el tamaño de muestra total es el mismo que quiere utilizar el estadístico A.





Solución

- a) En un muestreo estratificado sin reemplazamiento tenemos que la varianza del estimador de la proporción es:

$$V_{st}(\hat{P}_{st}) = \sum_{h=1}^L W_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{N_h \cdot P_h \cdot Q_h}{(N_h - 1)n_h} \approx \sum_{h=1}^L W_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{P_h \cdot Q_h}{n_h}$$

haciendo la suposición de que $N_h - 1 \approx N_h$, puesto que el número de árboles de la población en cada estrato es suficientemente grande.

Desarrollando esta expresión, y teniendo en cuenta que al utilizar el estadístico A la afijación proporcional se cumple que $n_h = n \cdot W_h = n \cdot \frac{N_h}{N}$ obtenemos lo siguiente:

$$\begin{aligned} V_{st}(\hat{P}_{st}) &\approx \sum_{h=1}^L W_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{P_h \cdot Q_h}{n_h} = \sum_{h=1}^L \frac{W_h^2 \cdot P_h \cdot Q_h}{n_h} + \sum_{h=1}^L \frac{W_h^2 \cdot P_h \cdot Q_h}{N_h} = \\ &= \frac{0,5^2 \cdot 0,7 \cdot 0,3}{n_1} + \frac{0,3^2 \cdot 0,4 \cdot 0,6}{n_2} + \frac{0,2^2 \cdot 0,15 \cdot 0,85}{n_3} - \frac{0,5^2 \cdot 0,7 \cdot 0,3}{10000} - \\ &\quad - \frac{0,3^2 \cdot 0,4 \cdot 0,6}{6000} - \frac{0,2^2 \cdot 0,15 \cdot 0,85}{4000} = \\ &= \frac{0,5^2 \cdot 0,7 \cdot 0,3}{n \cdot 0,5} + \frac{0,3^2 \cdot 0,4 \cdot 0,6}{n \cdot 0,3} + \frac{0,2^2 \cdot 0,15 \cdot 0,85}{n \cdot 0,2} - 0,000010125 = \\ &= \frac{0,105}{n} + \frac{0,072}{n} + \frac{0,0255}{n} - 0,000010125 = \frac{0,2025}{n} - 0,000010125 \end{aligned}$$

Con error máximo $e = 0,01$ lo que haremos es despejar n de la expresión $e = \sqrt{V_{st}(\hat{P}_{st})}$:

$$e^2 = 0,01^2 = \frac{0,2025}{n} - 0,000010125$$

Luego $\frac{0,2025}{n} = 0,000110125$, es decir, que $n = 1838,82 \approx 1839$

- b) Sustituyendo el tamaño de muestra anteriormente calculado en la expresión de la afijación proporcional obtenemos la distribución de la muestra en cada estrato:

$$\begin{aligned} n_1 &= 1839 \times \frac{10000}{20000} \approx 919,5 \\ n_2 &= 1839 \times \frac{6000}{20000} \approx 551,7 \\ n_3 &= 1839 \times \frac{4000}{20000} \approx 367,8 \end{aligned}$$

- c) El estadístico B utiliza un muestreo estratificado con reemplazamiento y afijación óptima. Podríamos calcular el tamaño de muestra para cometer un error



igual al del apartado a) y eso nos daría un tamaño de muestra diferente a dicho apartado lógicamente, pero como quiere usar el tamaño de muestra total calculado anteriormente, teniendo en cuenta los costes en cada estrato, que son: $C_1 = 4\text{£}$, $C_2 = 4\text{£}$, $C_3 = 9\text{£}$ y dado que $n \approx 1839$ es conocido que:

$$n_h = n \times \frac{N_h \sqrt{\frac{P_h(1-P_h)}{C_h}}}{\sum_h N_h \sqrt{\frac{P_h(1-P_h)}{C_h}}}$$

Cáculamos los numeradores de esa expresión:

$$N_1 \sqrt{\frac{P_1(1-P_1)}{C_1}} = 10000 \times \sqrt{\frac{0,7(1-0,7)}{4}} = 2291,3$$

$$N_2 \sqrt{\frac{P_2(1-P_2)}{C_2}} = 6000 \times \sqrt{\frac{0,4(1-0,4)}{4}} = 1469,7$$

$$N_3 \sqrt{\frac{P_3(1-P_3)}{C_3}} = 4000 \times \sqrt{\frac{0,15(1-0,15)}{9}} = 301,5$$

Sumando esas tres cantidades tenemos que el denominador vale 4062.5, por lo que los valores de la muestra en cada estrato son:

$$n_1 = 1839 \times \frac{2291,3}{4062,5} \approx 1037,22$$

$$n_2 = 1839 \times \frac{1469,7}{4062,5} \approx 665,3$$

$$n_3 = 1839 \times \frac{301,5}{4062,5} \approx 136,5$$



2.2. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte Común.

Cuestión 1. Se considera una población finita de tamaño N , compuesta por unidades ordenadas de acuerdo con una variable de interés definida por:

$$y_k = k, \quad \text{para } k = 1, 2, \dots, N.$$

donde los parámetros poblacionales de esta variable son:

- Media: $\mu_y = \frac{N+1}{2}$,
- Varianza: $\sigma_y^2 = \frac{(N-1)(N+1)}{12}$.

Además, se supone que $N = n \cdot a$, con $n, a \in \mathbb{N}$.

El objetivo es analizar la varianza del estimador de Horvitz-Thompson de la media poblacional bajo distintos esquemas de muestreo. Para ello, se plantea:

- (a) Suponiendo que se selecciona una muestra aleatoria simple sin reemplazamiento de tamaño n , deduzca la expresión de la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (b) La población se divide en n estratos consecutivos de tamaño a , de acuerdo al orden según la variable de interés, extrayéndose una unidad aleatoriamente por estrato. Deduzca la varianza del estimador de Horvitz-Thompson de la media poblacional en el muestreo estratificado descrito, sabiendo que la varianza poblacional dentro de cada estrato es $\sigma_i^2 = \frac{a^2-1}{12}, \forall i = 1, \dots, n$.
- (c) Suponer que se selecciona una muestra de tamaño n mediante muestreo sistemático con probabilidades iguales aplicado sobre la población ordenada de acuerdo a la variable de interés. Obtenga la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (d) Compare los tres planes de muestreo y justifique cuál resulta más adecuado en la situación planteada.

Solución.

- (a) Para deducir la expresión de la varianza del estimador de Horvitz-Thompson de la media poblacional bajo muestreo aleatorio simple sin reemplazamiento



(MAS), escribimos su expresión para la media poblacional μ_y que viene dada por:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k \in S} \frac{y_k}{\pi_k},$$

donde S es la muestra seleccionada, y_k es el valor de la variable de interés para la unidad k , y π_k es la probabilidad de inclusión de la unidad k en la muestra.

En el caso de MAS sin reemplazamiento de tamaño n , la probabilidad de inclusión π_k es la misma para todas las unidades y está dada por:

$$\pi_k = \frac{n}{N}, \quad \forall k = 1, 2, \dots, N.$$

Sustituyendo π_k en el estimador, obtenemos:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k \in S} \frac{y_k}{\frac{n}{N}} = \frac{1}{n} \sum_{k \in S} y_k = \bar{y},$$

donde $\bar{y}$ es la media muestral.

La varianza de $\bar{y}$ en MAS sin reemplazamiento está dada por (ver nota 1 al final del ejercicio):

$$\text{Var}(\bar{y}) = (1 - f) \frac{S_y^2}{n},$$

donde: $f = \frac{n}{N}$ es la fracción de muestreo y $S_y^2 = \frac{1}{N-1} \sum_{k=1}^N (y_k - \mu_y)^2$ es la cuasivarianza poblacional.

Dado que $\sigma_y^2 = \frac{1}{N} \sum_{k=1}^N (y_k - \mu_y)^2$, se cumple:

$$S_y^2 = \frac{N}{N-1} \sigma_y^2.$$

Sustituyendo σ_y^2 con el valor que nos proporciona el enunciado obtenemos:

$$S_y^2 = \frac{N}{N-1} \cdot \frac{(N-1)(N+1)}{12} = \frac{N(N+1)}{12}.$$

Por tanto, la varianza del estimador queda:

$$\text{Var}(\hat{\mu}_{HT}) = \left(1 - \frac{n}{N}\right) \frac{S_y^2}{n} = \left(\frac{N-n}{N}\right) \frac{N(N+1)}{12n}.$$

Simplificando:

$$\text{Var}_{MAS}(\hat{\mu}_{HT}) = \frac{(N-n)(N+1)}{12n} = \frac{(a-1)(na+1)}{12}.$$





- (b) Tenemos una población de tamaño $N = n \cdot a$, ordenada según la variable de interés $y_k = k$ para $k = 1, \dots, N$. La población se divide en n estratos consecutivos de tamaño a cada uno, y se selecciona 1 unidad por estrato de forma aleatoria sin reemplazamiento, lo que hace que la afijación sea igual y constante en todos los estratos.

Sabemos que la varianza intra-estrato vale: $\sigma_i^2 = \frac{a^2-1}{12}$, para todo estrato i , y que las unidades del estrato i son:

$$(i-1)a + 1, \quad (i-1)a + 2, \quad \dots, \quad ia.$$

Esta es una secuencia aritmética cuyo primer término es $y_{\min} = (i-1)a + 1$, su último término vale $y_{\max} = ia$, y hay a términos o elementos en el estrato. En muestreo estratificado, el estimador de Horvitz-Thompson para la media poblacional es:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{i=1}^n \frac{y_i}{\pi_i},$$

donde: y_i es el valor de la unidad seleccionada en el estrato i y $\pi_i = \frac{1}{a}$ (probabilidad de inclusión dentro de cada estrato, al seleccionar 1 unidad de a posibles)

Como seleccionamos 1 unidad por estrato:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{i=1}^n \frac{y_i}{1/a} = \frac{1}{na} \sum_{i=1}^n a y_i = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}_{\text{str}},$$

que es simplemente la media de las observaciones muestrales.

Es conocido que la varianza del estimador de la media tiene la expresión, teniendo en cuenta que en este caso tenemos n estratos con a elementos cada estrato, puesto que $N = n \cdot a$, $N_h = a$, $n_h = 1 \forall h = 1, \dots, n$, y la afijación es proporcional al elegir un elemento en cada estrato, con lo que $W_h = \frac{a}{N} = \frac{1}{n}$:

$$\begin{aligned} \text{Var}(\hat{\mu}_{HT}) &= \text{Var}(\bar{y}_{\text{str}}) = \sum_{h=1}^n W_h^2 \frac{N_h - n_h}{N_h - 1} \frac{\sigma_h^2}{n_h} = \frac{1}{n^2} \sum_{i=1}^n \frac{a-1}{a-1} \frac{a^2-1}{12} = \frac{n(a^2-1)}{12n^2} = \\ &= \frac{a^2-1}{12n} = \frac{(N-n)(N+n)}{12n^3} \end{aligned}$$

- (c) En este tipo de muestreo se selecciona un número aleatorio $r \in \{1, 2, \dots, a\}$ (que se conoce como el punto de arranque) para posteriormente seleccionar las unidades:



$$r, r + a, r + 2a, \dots, r + (n - 1)a$$

De esta forma, se obtiene una muestra de tamaño n , y se recorre la población completa en saltos de tamaño a .

Como la población está perfectamente ordenada por la variable $y_k = k$, el muestreo sistemático toma una unidad de cada bloque de tamaño a , pero en posiciones fijas dependiendo del valor inicial r .

En general, el estimador de Horvitz-Thompson de la media poblacional tiene la expresión

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k}$$

En el caso del muestreo sistemático con probabilidades iguales (todos los elementos tienen la misma probabilidad de inclusión $\pi_k = \frac{n}{N} = \frac{1}{a}$) dicha expresión queda de la siguiente manera:

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k \in s} \frac{y_k}{1/a} = \frac{1}{N} \sum_{k \in s} a y_k = \frac{a}{N} \sum_{k \in s} y_k$$

Pero como la muestra tiene tamaño n y $N = n \cdot a$, entonces el estimador de Horvitz-Thompson de la media en este caso es simplemente la media muestral:

$$\hat{\mu}_{HT} = \frac{a}{n \cdot a} \sum_{k \in s} y_k = \frac{1}{n} \sum_{k \in s} y_k = \bar{y}_s = \frac{1}{n} \sum_{j=0}^{n-1} y_{r+ja}$$

Vamos a calcular ahora la varianza del estimador $\hat{\mu}_{HT}$ obtenido con muestreo sistemático. En este muestreo, al trabajar con un arranque aleatorio, hay a posibles muestras sistemáticas, una para cada valor posible del punto de arranque $r \in \{1, 2, \dots, a\}$.

Para cada valor de r , la muestra está formada por las unidades:

$$\{r, r + a, r + 2a, \dots, r + (n - 1)a\}$$

Entonces, la media de esta muestra es:

$$\bar{y}_{(r)} = \frac{1}{n} \sum_{j=0}^{n-1} y_{r+ja} = \frac{1}{n} \sum_{j=0}^{n-1} (r + ja)$$





Resolvemos esa suma:

$$\bar{y}_{(r)} = \frac{1}{n} \left(nr + a \sum_{j=0}^{n-1} j \right) = \frac{1}{n} \left(nr + a \cdot \frac{(n-1)n}{2} \right) = r + \frac{a(n-1)}{2}$$

A esta expresión se puede llegar también haciendo uso del estimador genérico del total en muestreo sistemático, que es insesgado incluso en el caso en que N no es múltiplo de a , es decir, $N = n \cdot a + c$ con c positivo, cuya expresión es $\hat{X}_{\text{sis}} = a \cdot \sum_{r=1}^n Y_r$. (ver nota 2, al final del ejercicio)

La única parte aleatoria del estimador es el valor de r , ya que los demás términos dependen de r pero son constantes.

Entonces:

$$\hat{\mu}_{HT} = \bar{y}_{(r)} = r + \frac{a(n-1)}{2}$$

Como $\frac{a(n-1)}{2}$ es constante, la varianza del estimador es la varianza de r :

$$\text{Var}(\hat{\mu}_{HT}) = \text{Var}(r)$$

Recordando que $r \sim \text{Uniforme}(1, 2, \dots, a)$, puede comprobarse fácilmente (ver nota 3 al final del ejercicio) que su varianza es:

$$\text{Var}(r) = \frac{a^2 - 1}{12}$$

por lo tanto finalmente llegamos a que:

$$\text{Var}(\hat{\mu}_{HT}^{\text{sis}}) = \frac{a^2 - 1}{12}$$

Este es el valor de la varianza del estimador de la media poblacional bajo muestreo sistemático con arranque aleatorio.

Como $N = na$ la varianza queda:

$$\text{Var}(\hat{\mu}_{HT}^{\text{sis}}) = \frac{(N+n)(N-n)}{12n^2}$$



- (d) Las tres varianzas obtenidas, todas en función de n y a , con $N = a \cdot n$ son las siguientes: Muestreo aleatorio simple (MAS):

$$\text{Var}(\hat{\mu}_{HT}^{\text{MAS}}) = \frac{(a-1)(na+1)}{12}$$

Muestreo estratificado (una unidad por estrato):

$$\text{Var}(\hat{\mu}_{HT}^{\text{estratificado}}) = \frac{a^2 - 1}{12n}$$

Muestreo sistemático (con arranque aleatorio, resultado del apartado c):

$$\text{Var}(\hat{\mu}_{HT}^{\text{sistemático}}) = \frac{a^2 - 1}{12}$$

En el cuadro 2.1 vamos a mostrar unos gráficos, para diferentes valores del tamaño poblacional que nos ayuden a ver cómo se comporta cada tipo de muestreo respecto a la varianza. Representamos tanto en la columna de la izquierda como en la primera fila, en el eje de abscisas el tamaño de muestra, con la particularidad en el último gráfico de que en ordenadas hemos calculado el logaritmo de la varianza, porque se aprecia mucho mejor la relación entre las varianzas. Por otra parte, en la columna derecha, en el segundo gráfico dibujamos el valor de a en abscisas frente a la varianza de los estimadores y en el tercero fijamos un valor de a , que es 20, y movemos el valor de n para ver cómo se comporta el logaritmo de la varianza. Como en el sistemático, dado un a fijo, la varianza no depende de n , por eso sale una línea en medio.



Capítulo 2. Exámenes 2024. Cuerpos de Estadística. Turno libre.

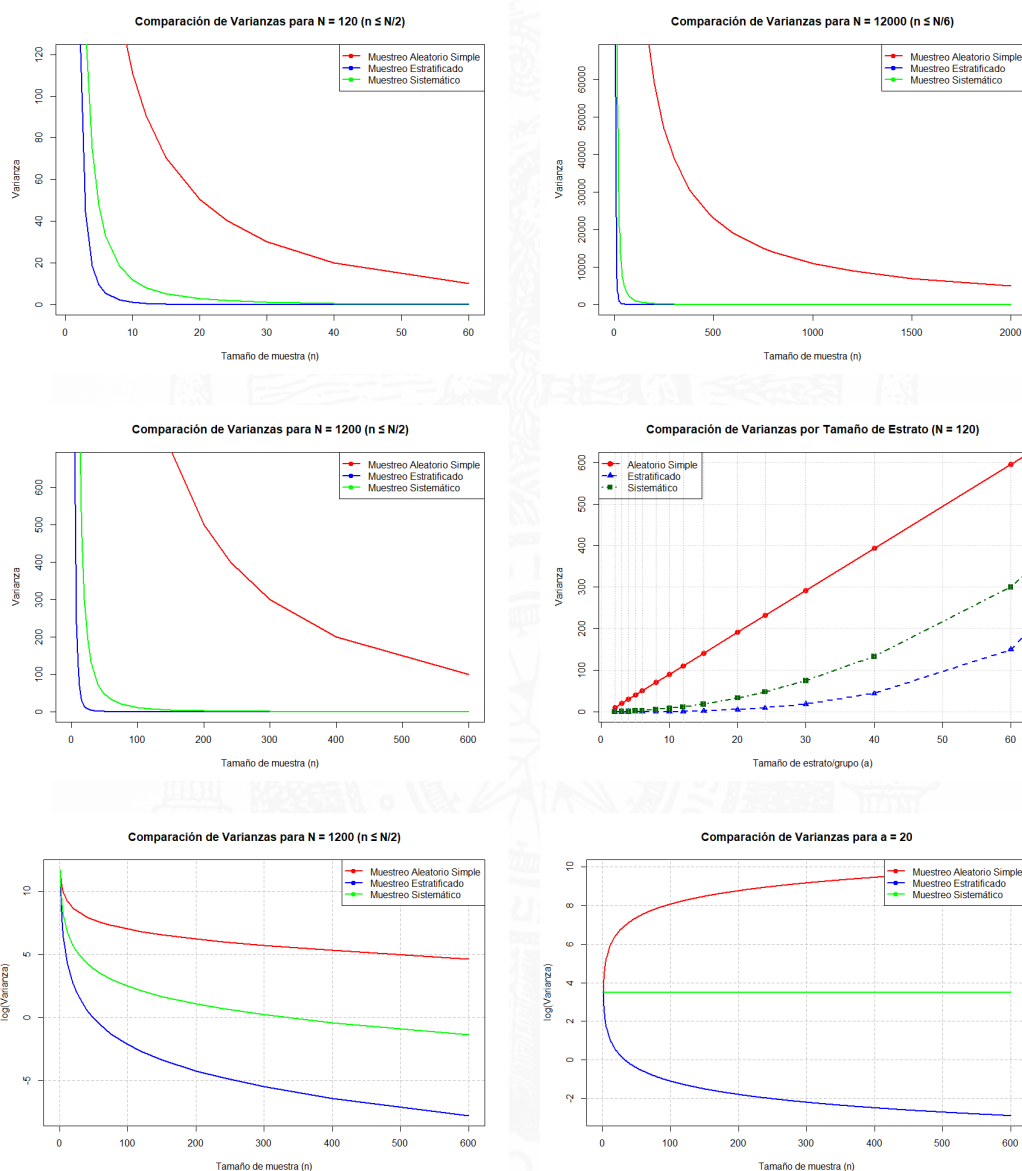


Tabla 2.1: Representaciones de la varianza o su logaritmo en función del tamaño de la muestra o del número de elementos por conglomerado para cada tipo de muestreo.



Puede observarse como el MAS es el menos eficiente, puesto que tiene una mayor varianza. El muestreo estratificado es el mejores resultados tiene, aunque no parecen muy diferentes a los del muestreo sistemático, a pesar de que la varianza del sistemático es la del estratificado dividida entre n .

En el estratificado, la variación dentro de cada grupo (de tamaño a) está controlada y podríamos decir que se reduce la variabilidad porque solo se selecciona una unidad de cada subgrupo homogéneo. En el sistemático, como se selecciona una unidad cada a posiciones, y la población está perfectamente ordenada como $y_k = k$, el patrón del muestreo aunque no puede reproducir lo que haría el muestreo estratificado, al ser los grupos homogéneos las varianzas no parecen tener un comportamiento muy diferente.

¿Nos podemos preguntar qué ocurre si cambia n o a ?

Conocemos que $N = a \cdot n$ (tamaño poblacional fijo)

Si $n \uparrow$ (más estratos, menor a) lo que ocurre es que en muestreo estratificado y en sistemático la varianza tiende a ser menor y se aprecia la diferencia entre ambos para n moderadamente pequeñas, aun cuando siempre es mucho menor la varianza del estratificado. En el muestreo aleatorio simple la varianza también disminuye, pero no tan rápido como en los otros diseños. Podríamos decir que el MAS no aprovecha la estructura ordenada de la población, por lo que su varianza es sistemáticamente mayor.

Para valores pequeños de a , es decir, con muchos estratos, los métodos estratificado y sistemático ofrecen una varianza muy baja, y a medida que a crece y hay menos estratos, las varianzas aumentan, pero el MAS siempre resulta el menos eficiente.

nota 1: La deducción de la fórmula de la varianza general del estimador de Horvitz-Thompson para el total en MAS sin reemplazamiento la detallamos a continuación.

Es conocido que el estimador de Horvitz-Thompson para el total poblacional es:

$$\hat{\tau}_{HT} = \sum_{k \in S} \frac{y_k}{\pi_k}$$

donde: S es la muestra seleccionada, y_k es el valor de la variable de interés para la unidad k y π_k es la probabilidad de inclusión de primer orden

Para calcular $\text{Var}(\hat{\tau}_{HT})$, partimos de la definición de varianza:





$$\text{Var}(\hat{\tau}_{HT}) = \mathbb{E}[\hat{\tau}_{HT}^2] - (\mathbb{E}[\hat{\tau}_{HT}])^2$$

Para calcular su esperanza, introducimos variables indicadoras I_k :

$$I_k = \begin{cases} 1 & \text{si la unidad } k \text{ está en la muestra } S, \\ 0 & \text{en caso contrario.} \end{cases}$$

El estimador puede reescribirse como:

$$\hat{\tau}_{HT} = \sum_{k=1}^N I_k \frac{y_k}{\pi_k},$$

donde ahora la suma es sobre todas las unidades poblacionales ($k = 1, \dots, N$).

Aplicando la linealidad de la esperanza:

$$\mathbb{E}[\hat{\tau}_{HT}] = \mathbb{E}\left[\sum_{k=1}^N I_k \frac{y_k}{\pi_k}\right] = \sum_{k=1}^N \frac{y_k}{\pi_k} \mathbb{E}[I_k].$$

Por definición de probabilidad de inclusión:

$$\mathbb{E}[I_k] = \pi_k.$$

Sustituyendo:

$$\mathbb{E}[\hat{\tau}_{HT}] = \sum_{k=1}^N \frac{y_k}{\pi_k} \cdot \pi_k = \sum_{k=1}^N y_k = \tau_y.$$

$$\begin{aligned} \mathbb{E}[\hat{\tau}_{HT}^2] &= \mathbb{E}\left[\left(\sum_{k \in S} \frac{y_k}{\pi_k}\right)^2\right] = \mathbb{E}\left[\sum_{k \in S} \sum_{l \in S} \frac{y_k y_l}{\pi_k \pi_l}\right] \\ &= \mathbb{E}\left[\sum_{k=1}^N \sum_{l=1}^N I_k I_l \frac{y_k y_l}{\pi_k \pi_l}\right] = \sum_{k=1}^N \sum_{l=1}^N \frac{y_k y_l}{\pi_k \pi_l} \mathbb{E}[I_k I_l] \end{aligned}$$

donde: $\mathbb{E}[I_k] = \pi_k$ y $\mathbb{E}[I_k I_l] = \pi_{kl}$ (probabilidad de inclusión de segundo orden)

Sustituyendo en la fórmula original:

$$\text{Var}(\hat{\tau}_{HT}) = \sum_{k=1}^N \sum_{l=1}^N \frac{y_k y_l}{\pi_k \pi_l} \pi_{kl} - \tau_y^2$$

Esta expresión puede reescribirse como:

$$\text{Var}(\hat{\tau}_{HT}) = \sum_{k=1}^N \sum_{l=1}^N (\pi_{kl} - \pi_k \pi_l) \frac{y_k y_l}{\pi_k \pi_l}$$



Para MAS sin reemplazamiento, las probabilidades de inclusión son:

$$\pi_k = \frac{n}{N}, \quad \pi_{kl} = \frac{n(n-1)}{N(N-1)} \quad \text{para } k \neq l.$$

Sustituyendo estas probabilidades en la expresión de la varianza:

$$\text{Var}(\hat{\tau}_{HT}) = \sum_{k=1}^N \left(\frac{n}{N} - \frac{n^2}{N^2} \right) \frac{y_k^2}{\left(\frac{n}{N} \right)^2} + \sum_{k \neq l} \left(\frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2} \right) \frac{y_k y_l}{\left(\frac{n}{N} \right)^2}.$$

Simplificando los términos:

$$\text{Var}(\hat{\tau}_{HT}) = \sum_{k=1}^N \left(\frac{N-n}{n} \right) y_k^2 + \sum_{k \neq l} \left(\frac{N(n-1) - n(N-1)}{n(N-1)} \right) y_k y_l.$$

Por lo tanto, la varianza queda:

$$\text{Var}(\hat{\tau}_{HT}) = \frac{N-n}{n} \sum_{k=1}^N y_k^2 - \frac{N-n}{n(N-1)} \sum_{k \neq l} y_k y_l.$$

Reescribimos las sumas en función de la media poblacional μ_y :

$$\sum_{k=1}^N y_k^2 = N\sigma_y^2 + N\mu_y^2, \quad \sum_{k \neq l} y_k y_l = \left(\sum_{k=1}^N y_k \right)^2 - \sum_{k=1}^N y_k^2 = N^2\mu_y^2 - N\sigma_y^2 - N\mu_y^2.$$

Sustituyendo en la varianza:

$$\text{Var}(\hat{\tau}_{HT}) = \frac{N-n}{n} (N\sigma_y^2 + N\mu_y^2) - \frac{N-n}{n(N-1)} (N^2\mu_y^2 - N\sigma_y^2 - N\mu_y^2).$$

Simplificando:

$$\text{Var}(\hat{\tau}_{HT}) = (N-n)N \left(\frac{\sigma_y^2 + \mu_y^2}{n} - \frac{N\mu_y^2 - \sigma_y^2 - \mu_y^2}{n(N-1)} \right).$$

Operando dentro del paréntesis:

$$\frac{\sigma_y^2 + \mu_y^2}{n} - \frac{(N-1)\mu_y^2 - \sigma_y^2}{n(N-1)} = \frac{(N-1)(\sigma_y^2 + \mu_y^2) - (N-1)\mu_y^2 + \sigma_y^2}{n(N-1)} = \frac{N\sigma_y^2}{n(N-1)}.$$

Por lo tanto:

$$\text{Var}(\hat{\tau}_{HT}) = (N-n)N \frac{N\sigma_y^2}{n(N-1)} = \frac{N^2(N-n)}{n(N-1)} \sigma_y^2.$$

Finalmente, la varianza del estimador de la media es:

$$\text{Var}(\hat{\mu}_{HT}) = \frac{1}{N^2} \text{Var}(\hat{\tau}_{HT}) = \frac{N-n}{n(N-1)} \sigma_y^2 = (1-f) \frac{S_y^2}{n}.$$





nota 2: En muestreo sistemático, en general, es $\hat{X}_{\text{sis}} = a \cdot \sum_{r=1}^n Y_r$. En este tipo de muestreo hay a muestras $s_1, \dots, s_a$ de n elementos cada una, siendo una de ellas dada por:

$$s = \{k : k = r + (j-1)a \leq N; j = 1, 2, \dots, n\} = s_r$$

La suma de los elementos de una de estas muestras es:

$$S_r = \sum_{r=1}^n Y_r = n \cdot r + a \cdot (0 + 1 + 2 + \dots + (n-1)) = n \cdot r + a \frac{(n-1)n}{2}$$

Luego $\hat{X}_{\text{sis}} = a \cdot n \cdot r + \frac{a^2 \cdot n(n-1)}{2}$, con lo que

$$\widehat{X}_{\text{sis}} = \frac{a \cdot r \cdot n}{N} + \frac{a^2 \cdot (n-1) \cdot n}{2N} = r + \frac{a(n-1)}{2}$$

nota 3: Para una distribución uniforme discreta que toma los valores $1, 2, \dots, a$, la varianza es:

$$\sigma^2 = \frac{a^2 - 1}{12}$$

La media de una variable aleatoria que siga esta distribución es $\mu = \frac{1+a}{2}$. El segundo momento ($\mathbb{E}[X^2]$) es

$$\mathbb{E}[X^2] = \frac{1}{a} \sum_{k=1}^a k^2 = \frac{(a+1)(2a+1)}{6}$$

Luego la varianza valdrá:

$$\begin{aligned} \sigma^2 &= \mathbb{E}[X^2] - \mu^2 = \\ &= \frac{(a+1)(2a+1)}{6} - \left(\frac{1+a}{2}\right)^2 = \frac{a^2 - 1}{12} \end{aligned}$$



Cuestión 2. Un investigador está analizando la duración (en horas) de cierto tipo de bombillas LED. Se asume que la duración de las bombillas sigue una distribución exponencial con parámetro λ . Se toma una muestra aleatoria simple de tamaño $n = 10$, con los siguientes datos (en horas):

$$\{88, 102, 96, 91, 110, 85, 99, 107, 95, 100\}$$

El fabricante afirma que el tiempo medio de vida es de al menos 100 horas.

- (a) Obtenga en el contexto de este problema, el valor del parámetro λ y qué significado tiene el mismo.
- (b) Formule el contraste de hipótesis para comprobar si los datos contradicen la afirmación del fabricante utilizando un contraste de razón de verosimilitudes.
- (c) Calcule la función de verosimilitud bajo la hipótesis nula y bajo la alternativa (en este caso usa el estimador máximo verosímil)
- (d) Calcule el estadístico de razón de verosimilitudes y determine si se rechaza H_0 al nivel $\alpha = 0,05$.
- (e) Interprete el resultado en el contexto del problema.

Solución.

- (a) En cuanto al valor del parámetro λ y su significado, sabemos que si $X \sim \text{Exp}(\lambda)$ es una distribución exponencial, el parámetro λ es la tasa de fallo (fallos por unidad de tiempo), y está relacionado con la media μ de la distribución, puesto que:

$$\mu = \frac{1}{\lambda}$$

Dado que el fabricante afirma que el tiempo medio de vida es de al menos 100 horas, bajo H_0 (hipótesis nula), tenemos:

$$\mu_0 = 100 \quad \Rightarrow \quad \lambda_0 = \frac{1}{100} = 0,01 \text{ (fallos por hora)}$$

- (b) Queremos contrastar la hipótesis del fabricante de que el tiempo medio de vida es de al menos 100 horas:

$$H_0 : \mu \geq 100 \quad (\text{equivalentemente, } \lambda \leq 0,01)$$

$$H_1 : \mu < 100 \quad (\text{equivalentemente, } \lambda > 0,01)$$

es decir, que nuestro contraste será:

$$\begin{cases} H_0 : \lambda \leq 0,01 \\ H_1 : \lambda > 0,01 \end{cases}$$





(c) Para una muestra aleatoria $x_1, \dots, x_n$, la función de verosimilitud es:

$$f_{\lambda}(x_1, \dots, x_n) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

Para calcular el valor de esta función en ambas hipótesis necesitamos conocer el estimador máximo verosímil, que es lo que vamos a calcular a continuación.

Tomando el logaritmo de la función de verosimilitud obtenemos:

$$L(x_1, \dots, x_n; \lambda) = \ln f_{\lambda}(x_1, \dots, x_n; \lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^n x_i$$

Y ahora derivamos respecto de λ e igualamos a cero:

$$\frac{\partial \ln f_{\lambda}(x_1, \dots, x_n)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

Despejando λ obtenemos que $\hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}$.

Tenemos que bajo H_0 ($\lambda \leq 0,01$), en la frontera, la función de verosimilitud vale:

$$f_{\lambda_0}(x_1, \dots, x_n) = (0,01)^{10} e^{-0,01 \times 973} \approx (0,01)^{10} e^{-9,73}$$

Como nos pide el enunciado, bajo H_1 usamos el estimador máximo verosímil y tendríamos $\lambda_{H_1} = \hat{\lambda}_{MV}$, y el EMV para λ es:

$$\hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^n x_i} = \frac{10}{973} \approx 0,010384$$

$$f_{\lambda_{H_1}}(x_1, \dots, x_n) = \left(\frac{10}{973} \right)^{10} e^{-10}$$

(d) El estadístico razón de verosimilitudes Λ se define como:

$$\Lambda(x_1, \dots, x_n) = \frac{\sup_{\lambda \leq \lambda_0} f_{\lambda}(x_1, \dots, x_n)}{\sup_{\lambda \in \mathbb{R}} f_{\lambda}(x_1, \dots, x_n)} = \frac{f_{\hat{\lambda}_{H_0}}(x_1, \dots, x_n)}{f_{\hat{\lambda}_{MV}}(x_1, \dots, x_n)}$$

Un test de razón de verosimilitudes para dichas hipótesis es aquel que tiene región crítica de la forma $C = \{\Lambda(x_1, \dots, x_n) < c\}$, es decir, que rechaza H_0 cuando el cociente Λ toma valores pequeños inferiores a un cierto valor crítico c .

Como f_{λ} (o su logaritmo) es una función creciente para $\lambda \leq \frac{n}{\sum_{i=1}^n x_i} = \hat{\lambda}_{MV}$ y decreciente para $\lambda > \frac{n}{\sum_{i=1}^n x_i} = \hat{\lambda}_{MV}$, el máximo bajo H_0 ($\lambda \leq 0,01 = \lambda_0$) se alcanza en:



- Si $\hat{\lambda}_{MV} \leq 0,01$, el EMV restringido es $\hat{\lambda}_{H_0} = \hat{\lambda}_{MV}$
- Si $\hat{\lambda}_{MV} > 0,01$, el EMV restringido es $\hat{\lambda}_{H_0} = 0,01$ (frontera de H_0)

En este segundo caso, el cociente quedaría de la siguiente forma:

$$\Lambda(x_1, \dots, x_n) = \frac{f_{\hat{\lambda}_{H_0}}(x_1, \dots, x_n)}{f_{\hat{\lambda}_{MV}}(x_1, \dots, x_n)} = \frac{0,01^n \cdot e^{-0,01 \sum_{i=1}^n x_i}}{\hat{\lambda}_{MV}^n \cdot e^{-\hat{\lambda}_{MV} \sum_{i=1}^n x_i}} = \left(\frac{0,01}{\hat{\lambda}_{MV}} \right)^n e^{\sum_{i=1}^n x_i (\hat{\lambda}_{MV} - 0,01)}$$

Con estas consideraciones, tenemos finalmente que la razón de verosimilitudes queda:

$$\Lambda(x_1, \dots, x_n) = \begin{cases} 1 & \text{si } \hat{\lambda}_{MV} \leq 0,01 \\ \left(\frac{0,01}{\hat{\lambda}_{MV}} \right)^n e^{\sum_{i=1}^n x_i (\hat{\lambda}_{MV} - 0,01)} & \text{si } \hat{\lambda}_{MV} > 0,01 \end{cases}$$

Además, es conocido que bajo H_0 , $-2 \ln \Lambda(x_1, \dots, x_n)$ sigue asintóticamente una distribución χ_1^2 :

$$-2 \ln \Lambda(x_1, \dots, x_n) = \begin{cases} 0 & \text{si } \hat{\lambda}_{MV} \leq 0,01 \\ -2 \sum_{i=1}^n x_i \left(\hat{\lambda}_{MV} - 0,01 - \hat{\lambda}_{MV} \ln \left(\frac{\hat{\lambda}_{MV}}{0,01} \right) \right) & \text{si } \hat{\lambda}_{MV} > 0,01 \end{cases}$$

siendo la regla de decisión, puesto que $\{\Lambda < c\}$ equivale a $\{-2 \ln \Lambda > c'\}$:

- Rechazar H_0 (evidencia de $\mu < 100$) si $-2 \ln \Lambda(x_1, \dots, x_n) > \chi_{1,\alpha}^2$
- Nivel de significación típico: $\alpha = 0,05$ ($\chi_{1,0,95}^2 \approx 3,841$)

En nuestro caso, con los datos de los que disponemos según el problema, teniendo en cuenta que $n = 10$ y que $\sum_{i=1}^{10} x_i = 973$ el estadístico quedaría:

$$\Lambda(x_1, \dots, x_n) = \frac{(0,01)^{10} e^{-9,73}}{\left(\frac{10}{973} \right)^{10} e^{-10}} = \left(\frac{0,01 \times 973}{10} \right)^{10} e^{0,27} \approx (0,973)^{10} e^{0,27}$$

Calculamos $-2 \ln \Lambda(x_1, \dots, x_n)$:

$$-2 \ln \Lambda(x_1, \dots, x_n) \approx -2(10 \ln 0,973 + 0,27) \approx -2(-0,2727 + 0,27) \approx 0,0074$$

El valor crítico para χ_1^2 con $\alpha = 0,05$ es 3,841. Como $0,0074 < 3,841$, no se rechaza H_0 .

- (e) Por tanto, la interpretación es que a la vista de los datos del problema no hay evidencia suficiente ($\alpha = 0,05$) para rechazar la afirmación del fabricante de que el tiempo medio de vida es de al menos 100 horas. Los datos son consistentes con H_0 .





Cuestión 3. .

- (a) Sea un sistema de representación binario de números enteros con **8 bits en complemento a dos**. Indicar detalladamente como realizar la operación: $45 - 84$ en este sistema, así como el resultado final.
- (b) Se quiere expresar el número entero **-315** en un sistema binario en complemento a dos, ¿cuál es el número **mínimo** de bits que se necesitan?
- (c) Se quiere representar el número **6.25** en un sistema binario normalizado de números **punto flotante de 32 bits**. El primer bit es el del signo, los siguientes 8 bits representan el exponente (con un sesgo de 127) y los restantes bits la mantisa. Indicar detalladamente los pasos para representarlo y el resultado final .

Solución.

- (a) El complemento a 2 es un sistema de representación numérica que permite trabajar con números negativos en binario. Sus características clave son:
 - El bit más significativo (MSB) indica el signo: 0 para positivo, 1 para negativo.
 - Para números positivos: coincide con la representación binaria normal.
 - Para números negativos: se obtiene complementando bit a bit el valor absoluto y sumando 1.

Este sistema es preferible a otros para realizar este tipo de cálculos debido a una serie de particularidades, que se podrían resumir en las siguientes:

1. Elimina la redundancia del cero, puesto que en complemento a 1, el número 0 tiene dos representaciones:
 - 0000 (+0)
 - 1111 (-0)
2. Permite implementaciones más simples en hardware, porque esta operación se puede implementar con un sumador estándar más inversores, sin lógica adicional. En complemento a 1, si hay un acarreo al sumar, se debe sumar 1 al resultado final, lo cuál añade un paso extra en el hardware, haciéndolo más lento y complejo.



3. Extiende el rango de números negativos representables, por ejemplo, con 4 bits el complemento a 1 iría de -7 a 7 (con dos ceros) y el complemento a 2 desde -8 hasta +7 (con un único cero, y por tanto más rango)
4. Unifica las operaciones de suma y resta, es decir, que la misma operación suma/resta funciona para positivos y negativos

Veamos cómo representar los números que se proponen en este apartado. El número 45 representado con 8 bits en binario sería:

$$45_{10} = 32 + 8 + 4 + 1 = 00101101_2$$

2^7 (128)	2^6 (64)	2^5 (32)	2^4 (16)	2^3 (8)	2^2 (4)	2^1 (2)	2^0 (1)
0	0	1	0	1	1	0	1

De la misma forma, la representación de 84 con 8 bits es:

$$84_{10} = 64 + 16 + 4 = 01010100_2$$

La resta $45 - 84$ puede realizarse como $45 + (-84)$. En complemento a 2:

1. Calculamos el complemento a 1 de 84 (inversión bit a bit):

$$01010100_2 \rightarrow 10101011_2$$

2. Sumamos 1 para obtener el complemento a 2 (representación de -84):

$$10101011_2 + 1 = 10101100_2$$

Ahora realizamos la suma binaria $45 + (-84)$:

$$\begin{array}{r} 00101101 \text{ (45)} \\ + 10101100 \text{ (-84)} \\ \hline 11011001 \text{ (-39)} \end{array}$$

Vamos a verificar e interpretar el resultado 11011001_2 :

- $\text{MSB} = 1 \Rightarrow$ número negativo
- Calculamos su complemento a 2 para hallar el valor absoluto:

$$11011001 \rightarrow 00100110 + 1 = 00100111_2 = 39_{10}$$

- Por tanto, $11011001_2 = -39_{10}$





- (b) Nos proponemos representar el número -315 en complemento a 2 en este apartado.

Para n bits, es conocido que el rango de números representable es:

$$-2^{n-1} \leq N \leq 2^{n-1} - 1$$

1. Probamos con 9 bits: $-256 \leq N \leq 255$ por lo tanto, como $315 > 255$ ese número de bits es insuficiente.
2. Probamos con 10 bits:

$$-512 \leq N \leq 511 \quad (\text{cumple } -315 \geq -512)$$

por lo tanto son necesarios 10 bits para representarlo.

Los pasos a seguir son los siguientes:

1. Valor absoluto en binario:

$$315_{10} = 0100111011_2$$

2. Complemento a 1:

$$1011000100_2$$

3. Complemento a 2 (sumamos 1):

$$-315 = 1011000101_2$$

- (c) Vamos a representar el número 6.25 en punto flotante de 32 bits mediante la estructura IEEE 754, la cuál se resume de la siguiente forma:

Signo (1 bit)	Exponente (8 bits)	Mantisa (23 bits)
---------------	--------------------	-------------------

Lo primero que es necesario hacer es convertir a binario ese número

$$6,25_{10} = 110,01_2$$

La explicación del paso anterior es la siguiente:

$$\begin{aligned} 6 &= 4 + 2 = 110_2 \\ 0,25 &= \frac{1}{4} = 0,01_2 \end{aligned}$$

El siguiente paso podríamos decir que es la normalización, en la cuál expresamos el número binario en notación científica binaria:

$$1,1001_2 \times 2^2$$



Esta normalización asegura que el bit más significativo sea 1(implícito).

A continuación analizamos los componentes de la estructura IEEE 754 con el caso que tenemos, que serían:

- **Signo:** 0 (positivo)
- **Exponente:** $2 + 127 = 129_{10} = 10000001_2$
- **Mantisa:** 10010000000000000000000 (23 bits)

El sesgo del exponente (127) permite representar exponentes negativos. En punto flotante, la mantisa almacena sólo la parte fraccionaria después del 1 implícito. Esto maximiza la precisión con el número de bits disponible.

Por último, la representación final del 6.25 sería la siguiente:

$$\underbrace{0}_{\text{signo}} \underbrace{10000001}_{\text{exponente}} \underbrace{10010000000000000000000}_{\text{mantisa}}$$

En hexadecimal:

$$0100\ 0000\ 1100\ 1000\ 0000\ 0000\ 0000\ 0000_2 = 0x40C80000$$



Cuestión 4. En un país P , se han registrado los datos de población por edad y de nacimientos por edad de la madre, que se proporcionan en la tabla. Las cifras están expresadas en miles de personas. A partir de esos datos, calcule los siguientes indicadores:

- (a) Tasa bruta de natalidad.
- (b) Ratio de masculinidad al nacimiento.
- (c) Tasa global de fecundidad.
- (d) Tasa específica de fecundidad para el grupo de edad “De 30 a 34 años”.
- (e) Índice sintético de fecundidad: indique dos formas de obtenerlo y calcúlelo utilizando la forma abreviada.
- (f) Tasa bruta de reproducción: indique dos formas de obtenerla y calcúlela utilizando la forma abreviada.

¿Qué información necesitaría para poder calcular la tasa neta de reproducción?

(Miles de personas)	Nacimientos 2023			Población 01/01/2023			Población 01/01/2024		
	Total	Hombres	Mujeres	Total	Hombres	Mujeres	Total	Hombres	Mujeres
Menos de 15	0	0	0	3.269	1.682	1.587	3.216	1.655	1.561
De 15 a 19	3	1	1	1.281	661	621	1.323	683	640
De 20 a 24	13	6	6	1.252	645	607	1.302	670	632
De 25 a 29	28	14	14	1.287	655	633	1.312	669	643
De 30 a 34	53	27	26	1.400	705	695	1.420	717	703
De 35 a 39	47	24	23	1.543	769	774	1.535	766	769
De 40 a 45	16	8	8	1.870	938	932	1.819	910	909
De 45 a 49	2	1	1	2.028	1.024	1.004	2.050	1.033	1.017
50 o más	0	0	0	10.115	4.706	5.409	10.337	4.813	5.524
De 15 a 49	160	82	78	10.659	5.395	5.264	10.758	5.447	5.312
Total	160	82	78	24.042	11.783	12.260	24.310	11.914	12.396



Solución.

- (a) La tasa bruta de natalidad(TBN) se calcula como:

$$\text{TBN} = \left(\frac{\text{Nacimientos totales}}{\text{Población total media}} \right) \times 1000$$

$$\text{Población media} = \frac{24,042 + 24,310}{2} = 24,176 \text{ miles}$$

$$\text{TBN} = \left(\frac{160}{24,176} \right) \times 1000 = 6,62 \text{ nacimientos por 1000 habitantes}$$

- (b) La ratio de masculinidad al nacimiento es:

$$\text{Ratio} = \left(\frac{\text{Nacimientos hombres}}{\text{Nacimientos mujeres}} \right) \times 100$$

$$\text{Ratio} = \left(\frac{82}{78} \right) \times 100 = 105,13 \text{ hombres por 100 mujeres}$$

Dicha ratio muestra un ligero predominio de nacimientos masculinos (105 hombres por 100 mujeres), lo que es normal en poblaciones humanas

- (c) La tasa global de fecundidad(TGF) se corresponde con la expresión:

$$\text{TGF} = \left(\frac{\text{Nacimientos}}{\text{Población media de Mujeres 15-49 años}} \right) \times 1000$$

$$\text{Población media de Mujeres 15-49 años} = \frac{5,264 + 5,312}{2} = 5,288 \text{ miles}$$

$$\text{TGF} = \left(\frac{160}{5,288} \right) \times 1000 = 30,26 \text{ nacimientos por 1000 mujeres}$$

La TGF considera sólo mujeres en edad fértil (15-49 años), dando una medida más precisa que la TBN.

- (d) La tasa específica de fecundidad del grupo de edada 30-34 años) vale:

$$\text{TEF}_{30-34} = \left(\frac{\text{Nacimientos 30-34}}{\text{Población media de mujeres de 30-34 años}} \right) \times 1000$$

$$\text{Población media de mujeres de 30-34 años} = \frac{695 + 703}{2} = 699 \text{ miles}$$

$$\text{TEF}_{30-34} = \left(\frac{53}{699} \right) \times 1000 = 75,82 \text{ nacimientos por 1000 mujeres}$$

El grupo de 30-34 años muestra la mayor fecundidad (75.82‰), reflejando posiblemente el retraso en la maternidad.





(e) El índice sintético de fecundidad se puede calcular de dos formas que describimos:

1. Suma de tasas específicas por grupo de edad multiplicado por 5 (duración del intervalo), es decir,

$$ISF = 5 \cdot \sum_x TEF_x$$

2. Forma abreviada: $ISF = TGF \times \text{duración periodo fértil (35 años)} / 1000$

Vamos a utilizar, como indica el enunciado la forma abreviada para el cálculo del ISF:

$$ISF = 30,26 \times \frac{35}{1000} = 1,06 \text{ hijos/as por mujer}$$

El valor obtenido de 1.06 está por debajo del nivel de reemplazo (2.1), indicando seguramente un posible decrecimiento poblacional a largo plazo.

(f) La tasa bruta de reproducción puede obtenerse de dos formas:

1. $TBR = ISF \times \text{proporción de niñas en nacimientos}$
2. $TBR = \text{Suma de tasas específicas de niñas por grupo de edad} \times 5 = 5 \cdot \sum_x TEF_{fem,x}$

El cálculo abreviado de la misma nos lleva a que:

$$TBR = 1,06 \times \left(\frac{78}{160} \right) = 0,52 \text{ hijas por mujer}$$

El valor 0.52 sugiere que cada generación de mujeres sería aproximadamente la mitad de la anterior.

Para calcular la tasa neta de reproducción se necesitaría:

- Tablas de mortalidad por edad para mujeres
- Probabilidades de supervivencia de las madres desde el nacimiento hasta la edad media de maternidad
- Distribución de nacimientos de niñas por edad de la madre

La expresión de la tasa neta es:

$$TNR = TBR \times p_{EMM}$$

siendo p_{EMM} la probabilidad de sobrevivir a la edad media a la maternidad.



Capítulo 2.2. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte Común.





Cuestión 5. La estimación más reciente del Producto Interior Bruto a precios de mercado (PIBpm) para el año 2023 de la Contabilidad Nacional Anual de España fue de 1.498 miles de millones €(mm€). Se dispone, además, de la siguiente información por ramas de actividad en las mismas unidades:

Agricultura, ganadería, silvicultura y pesca:	
Consumos Intermedios	38
Valor añadido bruto	37
Remuneración de los asalariados:	8
Otros impuestos netos sobre la producción:	-6
Industria:	
Consumos Intermedios	603
Remuneración de los asalariados:	98
Otros impuestos netos sobre la producción:	1
Construcción:	
Consumos Intermedios	122
Valor añadido bruto	81
Remuneración de los asalariados:	48
Otros impuestos netos sobre la producción:	3
Servicios:	
Consumos Intermedios	679
Valor añadido bruto	1.029
Remuneración de los asalariados:	562
Otros impuestos netos sobre la producción:	16
Otra información contable disponible:	
Formación bruta de capital:	315
Importaciones de bienes y servicios:	511
Saldo de intercambios exteriores de bienes y servicios	-59
Excedente de explotación bruto/Renta mixta bruta de la economía	639

Se pide:

- Calcular el valor añadido bruto de la industria.
- Elaborar la cuenta de bienes y servicios de la economía simplificada para ese año.
- Sabiendo que en 2023 la tasa de variación interanual del índice de volumen encadenado del PIBpm fue del 2,7% y que el empleo equivalente a tiempo completo experimentó un aumento del 3,27%. ¿Qué se podría decir del comportamiento de la productividad en ese año?



Solución.

- (a) Para calcular el valor añadido bruto de la industria la idea es, teniendo el PIB como un dato del problema, y teniendo en cuenta que se cumple:

$$\text{PIB} = \sum_{\text{ramas}} \text{VAB}(\text{rama}) + (D.21 - D.31)$$

como tenemos los valores añadidos brutos de todas las ramas menos la que nos piden, si somos capaces de calcular los impuestos netos sobre los productos, despejando estaría resuelto el apartado.

Dichos impuestos netos sobre productos los podemos calcular utilizando la cuenta de explotación. En el siguiente cuadro, en gris están datos que el problema nos da directamente (que son el PIB y el EBE/RMB) y en naranja están los que calculamos a partir de los que nos dan, de la siguiente forma:

La remuneración de asalariados de las 4 ramas es:

$$\begin{aligned} \text{RA}(\text{total}) &= \text{RA}(\text{agricultura}) + \text{RA}(\text{industria}) + \text{RA}(\text{construcción}) + \text{RA}(\text{servicios}) = \\ &= 8 + 98 + 48 + 562 = 716 \end{aligned}$$

Los otros impuestos netos sobre producción de las 4 ramas es:

$$D.29 - D.39 = \sum_{\text{ramas}} (D.29 - D.39)(\text{rama}) = -6 + 1 + 3 + 16 = 14$$

II. 1.1 Cuenta de explotación			
Empleos	Código	Operaciones y otros flujos y saldos contables	
S.1			S.1
Total de la economía			Total de la economía
	B.1b/B.1*b	Producto interior bruto	1.498
716	D.1	Remuneración de los asalariados(Interior)	
x	D.21-D.31	Impuestos netos sobre los productos	
14	D.29 -D.39	Otros impuestos netos sobre la producción	
639	B.2b + B.3b	Excedente de explotación bruto + Renta mixta bruta	

Despejando por tanto la incógnita x, que son los impuestos netos sobre los productos tenemos que

$$x = D.21 - D.31 = 1498 - 716 - 14 - 639 = 129$$



Utilizando por último la primera ecuación que comentábamos podemos sustituir los valores conocidos obteniendo

$$1498 = 37 + \text{VAB(industria)} + 81 + 1029 + 129$$

por lo que el valor que se nos pide será:

$$\text{VAB(industria)} = 1498 - 1029 - 37 - 81 - 129 = 222$$

- (b) Para elaborar la cuenta de bienes y servicios tenemos que calcular algunos agregados con los datos que nos da el problema, como son la producción, los consumos intermedios, las importaciones y las exportaciones. La producción, teniendo en cuenta que para cada rama es la suma del valor añadido bruto y los consumos intermedios, vale

$$\begin{aligned} \text{Producción} &= \sum_{\text{ramas}} \text{Producción(rama)} = (37 + 38) + (603 + 222) + \\ &+ (122 + 81) + (679 + 1029) = 2811 \end{aligned}$$

Los consumos intermedios totales son

$$\text{Cons.Inter} = \sum_{\text{ramas}} \text{CI(rama)} = 38 + 603 + 122 + 679 = 1442$$

Además, como el saldo de intercambios exteriores, que son las exportaciones menos las importaciones tiene un valor negativo de -59, al ser las importaciones 511, entonces es claro que las exportaciones valdrán $511 - 59 = 452$. En este punto, podemos ver la cuenta en el siguiente cuadro:



Cuenta de bienes y servicios			
Recursos	Operaciones		Empleos
2.023			2.023
2.811	P.1	Producción	
	P.2	Consumos intermedios	1.442
129	D.21 - D.31	Impuestos netos sobre los productos	
	P.3/	Gasto en consumo final	x
	P.5	Formación bruta de capital	315
	P.6	Exportaciones de bienes y servicios	452
511	P.7	Importaciones de bienes y servicios	
3.451			3.451

Como los recursos y los empleos deben cuadrar, en la columna de los recursos la suma total es 3451. No conocemos el agregado gasto en consumo final, pero por diferencia se puede calcular y vale

$$\text{GCF} = 3451 - 1442 - 315 - 452 = 1242$$

- (c) La productividad, de una manera genérica en este contexto, se podría definir como el cociente entre el PIB real(volumen) y el empleo equivalente a tiempo completo, es decir:

$$\text{Productividad} = \frac{\text{PIB real}}{\text{Empleo}}$$

Por lo tanto(ver nota al final del ejercicio) la tasa de variación de la productividad se puede aproximar como:

$$\Delta \% \text{Productividad} \approx \Delta \% \text{PIB} - \Delta \% \text{Empleo}$$

Con los datos proporcionados tenemos lo siguiente:

- Tasa de variación interanual del PIB real (volumen): $\Delta \% \text{PIB} = 2,7 \%$
- Tasa de variación del empleo equivalente a tiempo completo: $\Delta \% \text{Empleo} = 3,27 \%$

por lo tanto, la tasa de variación será:

$$\Delta \% \text{Productividad} \approx 2,7 \% - 3,27 \% = -0,57 \%$$

que se puede interpretar como que la productividad derivada del trabajo disminuyó un 0,57 % en 2023, es decir, que el empleo creció a mayor ritmo que



el PIB (3,27 % vs 2,7 %). Hay que apuntar que esta aproximación solo funciona cuando las variaciones del PIB y del empleo son pequeñas, y a medida que estos cambios tienen una entidad considerable este cálculo no funciona bien y no quedaría más remedio que calcular explícitamente dicho incremento porcentual.

Este valor negativo en el incremento porcentual de la productividad pudo ser debido a diversas causas, como:

- Incorporación de trabajadores menos productivos (ej. en sectores con menor valor añadido)
- Reducción de horas trabajadas sin disminución proporcional en producción
- Ineficiencias productivas o cambios en la estructura sectorial

Podríamos decir, a la vista de estos datos que en 2023, España experimentó un deterioro de la productividad del trabajo ($-0,57\%$), reflejando que el crecimiento del empleo superó al del PIB real. Esto sugiere un uso menos eficiente del factor trabajo respecto a 2022.

nota: La tasa de variación relativa(o tasa de crecimiento instantánea) de una variable $x(t)$ se define como:

$$\text{Tasa de variación relativa} = \frac{\text{Cambio porcentual en } x}{\text{Cambio en el tiempo}} = \frac{x'(t)}{x(t)}$$

donde:

- $x'(t) = \frac{dx}{dt}$ es la derivada de x respecto al tiempo (su velocidad de cambio).
- $x(t)$ es el valor de la variable en el instante t .

Esta definición tiene relación con la derivada logarítmica de $x(t)$, que se define como $\frac{d}{dt}(\ln x(t))$, puesto que aplicando la regla de la cadena al derivar $\ln x(t)$, obtenemos:

$$\frac{d}{dt}(\ln x(t)) = \frac{1}{x(t)} \cdot \frac{dx}{dt} = \frac{x'(t)}{x(t)}$$

que es justamente tasa de variación relativa que definimos antes.

Nosotros tenemos un cociente, que es el PIB real entre el empleo, que lo hemos llamado productividad, pero en general sea $z(t) = \frac{x(t)}{y(t)}$, donde $x(t)$ y $y(t)$ son



funciones diferenciables del tiempo. Queremos comprobar que la tasa de variación relativa de z es igual a la diferencia de las tasas de variación de x e y . Para ello aplicamos logaritmo natural a ambos lados para linealizar la relación:

$$\ln z(t) = \ln x(t) - \ln y(t)$$

Derivando respecto al tiempo t (regla de la cadena) queda:

$$\frac{d}{dt} \ln z(t) = \frac{d}{dt} \ln x(t) - \frac{d}{dt} \ln y(t)$$

Lo que equivale a:

$$\frac{z'(t)}{z(t)} = \frac{x'(t)}{x(t)} - \frac{y'(t)}{y(t)}$$

que es precisamente la diferencia entre las tasas de variación relativa de x e y . Para cambios pequeños Δx y Δy , se cumple aproximadamente:

$$\frac{\Delta z}{z} \approx \frac{\Delta x}{x} - \frac{\Delta y}{y}$$

Por lo tanto hemos visto que la tasa de variación de un cociente es exactamente la diferencia de las tasas de variación del numerador y denominador, tanto en términos instantáneos (derivadas) como para pequeñas variaciones discretas.



Cuestión 6. Las autoridades municipales desean conocer la utilización de un nuevo tramo del carril bici.

Para ello, se empleó un muestreo por conglomerados sin reemplazamiento y con probabilidades iguales. En concreto, se seleccionaron aleatoriamente 4 semanas del total de 52 que componen el año. Durante cada una de las semanas seleccionadas, se registró diariamente el número de bicicletas que circularon por un punto específico del nuevo tramo del carril bici.

Los datos recogidos se presentan en la siguiente tabla:

Semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Total Semana
1	20	18	15	17	22	25	30	147
2	12	10	13	11	14	15	17	92
3	14	16	12	13	15	20	21	111
4	15	13	10	12	14	17	18	99

Se pide:

- Una estimación insesgada del total de usuarios anuales del nuevo tramo del carril bici y su error de muestreo relativo o coeficiente de variación. Calcule también un intervalo de confianza del 95 % para dicho total. Considerar $z_{\frac{\alpha}{2}} = 2$.
- Dichas autoridades estarían interesadas en estimar el número medio diario de usuarios del carril bici en un año futuro utilizando el mismo diseño muestral. Con el objetivo de garantizar que la varianza del estimador del número medio de usuarios diarios sea menor que uno, se pide determinar el número mínimo de semanas que deben seleccionarse.

Solución.

Este ejercicio podemos encuadrarlo dentro de la técnica de muestreo correspondiente a conglomerados sin submuestreo, siendo las semanas los conglomerados, puesto que se eligen 4 de 52 semanas, pero de cada semana se cogen todos los días, no unos pocos. Y cada conglomerado tiene el mismo número de elementos, que es 7, número de días de una semana.

En el anexo de este ejercicio se presentan las fórmulas y la notación utilizada en el desarrollo del mismo.

- En este tipo de muestreo por conglomerados sin reemplazamiento y con probabilidades iguales, la estimación insesgada del total poblacional se calcula como:

$$\hat{X} = \frac{N}{n} \sum_{i=1}^n X_i$$

siendo $N = 52$, que son el total de semanas (total de conglomerados en la población), $n = 4$, que son el número de semanas o conglomerados seleccionados en la muestra y x_i es el total de usuarios en la semana i .



De la tabla, los totales semanales son $X_1 = 147$, $X_2 = 92$, $X_3 = 111$, $X_4 = 99$.

$$\hat{X} = \frac{52}{4}(147 + 92 + 111 + 99) = 13 \times 449 = 5837$$

Para calcular el error de muestreo relativo o coeficiente de variación, CV) primero calculamos una estimación de la varianza del estimador del total:

$$\hat{V}(\hat{X}) = N^2 \left(1 - \frac{n}{N}\right) \frac{s_x^2}{n}$$

Donde s_x^2 en este caso se refiere a la cuasi-varianza de los totales semanales en la muestra:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{M\bar{x}})^2$$

Calculamos $\overline{M\bar{x}} = \frac{1}{4} \sum_{i=1}^4 X_i = \frac{449}{4} = 112,25$, puesto que la suma de los totales de las 4 semanas es $147 + 92 + 111 + 99 = 449$, y así:

$$\begin{aligned} s_x^2 &= \frac{1}{3} [(147 - 112,25)^2 + (92 - 112,25)^2 + (111 - 112,25)^2 + (99 - 112,25)^2] = \\ &= \frac{1}{3} [1206,25 + 410,0625 + 1,5625 + 175,5625] = \frac{1793,4375}{3} \approx 597,8125 \end{aligned}$$

Luego:

$$\hat{V}(\hat{X}) = 52^2 \left(1 - \frac{4}{52}\right) \frac{597,8125}{4} = 2704 \times 0,9231 \times 149,4531 \approx 373044$$

El error estándar es:

$$\sqrt{\hat{V}(\hat{X})} \approx \sqrt{373044} \approx 610,77$$

El coeficiente de variación es:

$$CV = \frac{\sqrt{\hat{V}(\hat{X})}}{\hat{X}} \times 100 = \frac{610,77}{5837} \times 100 \approx 10,46 \%$$

Un intervalo de confianza al 95

$$\hat{X} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\hat{V}(\hat{X})} = 5837 \pm 2 \times 610,77 = 5837 \pm 1221,54$$

El intervalo es:

$$(4615,46, 7058,54)$$





- (b) En este apartado nos piden el número mínimo de semanas para garantizar que la varianza del estimador del número medio diario sea menor que 1. Vamos a considerar que el número de días del año es 365 en el cálculo, aun cuando quizá sería más oportuno tomar que el número de días sea 364 para que se cumpla que $N \cdot \overline{M} = M$, siendo $N = 52$ y $\overline{M} = 7$ el número de días por semana. De cualquier forma se obtiene el mismo resultado y no es absolutamente relevante.

El número medio diario anual se estima como:

$$\hat{\mu} = \frac{\hat{X}}{365}$$

La varianza de $\hat{\mu}$ es:

$$V(\hat{\mu}) = \frac{V(\hat{X})}{365^2}$$

La condición que nos piden equivale a que $V(\hat{\mu}) < 1$, por lo que:

$$\frac{V(\hat{X})}{365^2} < 1 \implies V(\hat{X}) < 365^2 = 133225$$

En este punto tenemos que utilizar los datos del diseño muestral para considerar la estimación de esa varianza, de modo que utilizamos la expresión:

$$\hat{V}(\hat{X}) = N^2 \left(1 - \frac{n}{N}\right) \frac{s_x^2}{n}$$

Usamos $s_x^2 \approx 597,8125$ de la muestra inicial obtenida en el apartado anterior. Resolviendo para n :

$$52^2 \left(1 - \frac{n}{52}\right) \frac{597,8125}{n} < 133225$$

Simplificando:

$$2704 \left(\frac{52 - n}{52}\right) \frac{597,8125}{n} < 133225$$

$$\frac{2704 \times 597,8125 \times (52 - n)}{52n} < 133225$$

$$\frac{1616485 \times (52 - n)}{52n} < 133225$$

$$\frac{1616485 \times 52 - 1616485n}{52n} < 133225$$

$$\frac{84057220 - 1616485n}{52n} < 133225$$

$$84057220 - 1616485n < 6927700n$$



$$8405,220 < 8544185n$$

$$n > \frac{84057220}{8544185} \approx 9,84$$

Redondeando al entero superior, $n = 10$.

Anexo:

En este contexto tenemos que $\hat{X} = \sum_{i=1}^n \frac{X_i}{\pi_i} = \sum_{i=1}^n \frac{\sum_{j=1}^{\bar{M}} X_{ij}}{n/N} = \frac{N}{n} \sum_{i=1}^n \sum_{j=1}^{\bar{M}} X_{ij} = \frac{N}{n} \sum_{i=1}^n X_i$, siendo X_i el total de bicis en la semana i -ésima.

Sabemos que $\bar{X} = \frac{1}{N\bar{M}} \sum_{i=1}^N \sum_{j=1}^{\bar{M}} X_{ij}$ y que $\bar{X}_i = \frac{1}{\bar{M}} \sum_{j=1}^{\bar{M}} X_{ij} = \frac{X_i}{\bar{M}}$

luego $\bar{M} \cdot \bar{X}_i = \sum_{j=1}^{\bar{M}} X_{ij} = X_i$, y $\bar{M} \cdot \bar{X} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{\bar{M}} X_{ij} = \frac{1}{N} \sum_{i=1}^N X_i$

Además $\hat{X} = \frac{N\bar{M}}{n} \sum_{i=1}^n \bar{X}_i = N\bar{M} \cdot \bar{\bar{x}}$, con $\bar{\bar{x}} = \frac{1}{n} \sum_{i=1}^n \bar{X}_i$

y así

$V(\hat{X}) = N^2\bar{M}^2 V(\bar{\bar{x}}) = N^2\bar{M}^2(1-f) \frac{S_b^2}{n\bar{M}}$, siendo $S_b^2 = \frac{\sum_{i=1}^N \sum_{j=1}^{\bar{M}} (\bar{X}_i - \bar{X})^2}{N-1}$ la cuasi-varianza entre conglomerados, luego

$$\begin{aligned} V(\hat{X}) &= \frac{N^2\bar{M}^2(1-f)}{n\bar{M}} \cdot \frac{\sum_{i=1}^N \sum_{j=1}^{\bar{M}} (\bar{X}_i - \bar{X})^2}{N-1} = \frac{N^2(1-f)\bar{M}}{n} \frac{\sum_{i=1}^N \bar{M}(\bar{X}_i - \bar{X})^2}{N-1} \\ &= \frac{N^2(1-f)}{n} \frac{\sum_{i=1}^N (\bar{M}(\bar{X}_i - \bar{X}))^2}{N-1} = \frac{N^2(1-f)}{n} \frac{\sum_{i=1}^N (X_i - \frac{1}{N} \sum X_i)^2}{N-1} \end{aligned}$$

Un estimador insesgado de S_b^2 es $\hat{S}_b^2 = \frac{1}{n-1} \sum_{i=1}^n \sum_{j=1}^{\bar{M}} (\bar{X}_i - \bar{\bar{x}})^2$

Luego podemos considerar como aproximación de la varianza del total la expresión:

$$\begin{aligned} \hat{V}(\hat{X}) &= N^2\bar{M}^2(1-f) \frac{\hat{S}_b^2}{n\bar{M}} = \frac{N^2(1-f)}{n} \frac{\bar{M}}{n-1} \sum_{i=1}^n \bar{M} (\bar{X}_i - \bar{\bar{x}})^2 = \\ &= \frac{N^2(1-f)}{n} \frac{1}{n-1} \sum_{i=1}^n (\bar{M} (\bar{X}_i - \bar{\bar{x}}))^2 = \frac{N^2(1-f)}{n} \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{M}\bar{\bar{x}})^2 = \\ &= \frac{N^2(1-f)}{n} \frac{1}{n-1} \sum_{i=1}^n \left(X_i - \frac{1}{n} \sum_{i=1}^n X_i \right)^2 = \end{aligned}$$





$$= \frac{N^2(1-f)}{n} \cdot s_x^2$$

con s_x^2 la cuasi-varianza de los totales semanales en la muestra, puesto que

$$\bar{\bar{x}} = \frac{1}{n} \sum_{i=1}^n \bar{X}_i = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^{\bar{M}} X_{ij}}{\bar{M}}$$

y

$$\overline{M\bar{x}} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{\bar{M}} X_{ij} = \frac{1}{n} \sum_{i=1}^n X_i$$

Cuestión 7. Sean los siguientes modelos

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u \quad (\text{Modelo 1})$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u \quad (\text{Modelo 2})$$

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + u \quad (\text{Modelo 3})$$

donde las variables y, x_1, x_2, u son aleatorias tales que los modelos están correctamente definidos.

Se pide:

- Considere que $\mathbb{E}(u|x_1, x_2) = 0$, y obtenga el efecto parcial de x_2 sobre $E(y|x_1, x_2)$ para los modelos 1 y 2.
- Considere que $\mathbb{E}(u|x_1, x_2) = 0$, e interprete el significado de β_1 y β_2 en el modelo 3 (justifique la respuesta).
- Obtenga, para el Modelo 1, $\mathbb{E}(u)$ sabiendo que $\mathbb{E}(u|x_1, x_2) = k$, $k \in \mathbb{R}$
- Considere que $\mathbb{E}(u|x_1, x_2) = 0$, obtenga (si es posible con la información facilitada) y para el Modelo 1 las siguientes covarianzas: $\text{cov}(u, x_1)$, $\text{cov}(u, x_2^2)$.
- Explique el significado y alcance de $\mathbb{E}(u|x_1, x_2) = 0$ en el contexto de la modelización.

Solución.

- Calculamos el efecto parcial de x_2 sobre $\mathbb{E}(y|x_1, x_2)$ para el modelo 1, cuya expresión es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$



Dado que $\mathbb{E}(u|x_1, x_2) = 0$, la esperanza condicional es:

$$\mathbb{E}(y|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

El efecto parcial de x_2 sobre $\mathbb{E}(y|x_1, x_2)$ es la derivada parcial con respecto a x_2 :

$$\frac{\partial \mathbb{E}(y|x_1, x_2)}{\partial x_2} = \beta_2$$

Respecto al segundo modelo, de expresión

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u$$

la esperanza condicional es:

$$\mathbb{E}(y|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2$$

por lo que el efecto parcial de x_2 es:

$$\frac{\partial \mathbb{E}(y|x_1, x_2)}{\partial x_2} = \beta_2 + 2\beta_3 x_2$$

(b) La expresión del modelo 3 es

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + u$$

Dado que según el enunciado $\mathbb{E}(u|x_1, x_2) = 0$, la esperanza condicional es:

$$\mathbb{E}(\log(y)|x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2)$$

La interpretación de β_1 sería que ese coeficiente mide el cambio porcentual aproximado en y ante un aumento unitario en x_1 , manteniendo x_2 constante.

La justificación de esta interpretación se basa en calcular la derivada con respecto a x_1 :

$$\frac{\partial \mathbb{E}(\log(y)|x_1, x_2)}{\partial x_1} = \beta_1 \implies \frac{\Delta y}{y} \approx \beta_1 \Delta x_1.$$

Si $\Delta x_1 = 1$, esto implica un cambio porcentual aproximado de $\beta_1 \times 100\%$ en y . Esta aproximación es correcta siempre que el incremento de x_1 sea pequeño, porque si no ocurre esto para cambios de cualquier magnitud en x_1 , el cambio porcentual *exacto* en y se calcula como:

1. El valor inicial (antes del cambio):

$$\log(y_0) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) \implies y_0 = e^{\beta_0 + \beta_1 x_1 + \beta_2 \log(x_2)}.$$





2. El valor final (después de incrementar x_1 en 1 unidad):

$$\log(y_1) = \beta_0 + \beta_1(x_1 + 1) + \beta_2 \log(x_2) \implies y_1 = e^{\beta_0 + \beta_1(x_1 + 1) + \beta_2 \log(x_2)} = y_0 \cdot e^{\beta_1}.$$

Calculando el cambio porcentual exacto:

$$\frac{y_1 - y_0}{y_0} = e^{\beta_1} - 1.$$

Por lo tanto, el cambio porcentual exacto en y ante un aumento unitario en x_1 es:

$$(e^{\beta_1} - 1) \times 100\%.$$

De la misma forma, la interpretación de β_2 sería la elasticidad de y con respecto a x_2 , es decir, representa el cambio porcentual en y ante un aumento del 1 % en x_2 , manteniendo x_1 constante.

La justificación de nuevo, la obtenemos derivando con respecto a $\log(x_2)$:

$$\frac{\partial \mathbb{E}(\log(y)|x_1, x_2)}{\partial \log(x_2)} = \beta_2 \implies \frac{\Delta y}{y} \approx \beta_2 \cdot \frac{\Delta x_2}{x_2}.$$

(c) Vamos a calcular $\mathbb{E}(u)$ en el Modelo 1 dado que $\mathbb{E}(u|x_1, x_2) = k$.

La expresión del modelo 1 es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

Por la ley de esperanzas iteradas:

$$\mathbb{E}(u) = \mathbb{E}[\mathbb{E}(u|x_1, x_2)] = \mathbb{E}(k) = k$$

(d) Dado que $\mathbb{E}(u|x_1, x_2) = 0$, se cumple:

$$\text{cov}(u, x_1) = \mathbb{E}(ux_1) - \mathbb{E}(u)\mathbb{E}(x_1)$$

Como $\mathbb{E}(ux_1) = \mathbb{E}[\mathbb{E}(ux_1|x_1, x_2)] = \mathbb{E}[x_1 \mathbb{E}(u|x_1, x_2)] = \mathbb{E}(x_1 \cdot 0) = 0$ y $\mathbb{E}(u) = 0$, entonces:

$$\text{cov}(u, x_1) = 0$$

Análogamente:

$$\text{cov}(u, x_2^2) = \mathbb{E}(ux_2^2) - \mathbb{E}(u)\mathbb{E}(x_2^2)$$

$$\mathbb{E}(ux_2^2) = \mathbb{E}[\mathbb{E}(ux_2^2|x_1, x_2)] = \mathbb{E}[x_2^2 \mathbb{E}(u|x_1, x_2)] = 0$$

Por lo tanto:

$$\text{cov}(u, x_2^2) = 0$$



- (e) La condición $\mathbb{E}(u|x_1, x_2) = 0$ implica que el término de error u no está correlacionado con las variables explicativas x_1 y x_2 . Esto garantiza lo siguiente:
1. Insesgadez de los estimadores MCO: los coeficientes estimados $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ son insesgados para los parámetros poblacionales.
 2. Exogeneidad estricta: el error no contiene información útil para predecir y una vez controladas x_1 y x_2 .
 3. Interpretación causal: si x_1 y x_2 son asignados aleatoriamente (o se controlan adecuadamente), los efectos estimados pueden interpretarse como causales.

Esta condición es fundamental para la validez interna del modelo. Si no se cumple (por ejemplo, por omisión de variables relevantes o errores de medida), los estimadores serán sesgados y inconsistentes. Es decir, esta condición asegura que el modelo está correctamente especificado y que los estimadores son insesgados.



Cuestión 8. De las cuentas no financieras de los Sectores Institucionales de una Economía se conocen los siguientes datos correspondientes a un año T, en unidades monetarias:

Sector Hogares e ISFLSH:

- Producción: 364
- Consumos Intermedios: 94
- Gasto en consumo final: 720.
- Remuneración de los asalariados residentes: 588, de los que 3 son de empleadores no residentes.
- Renta mixta bruta: 137
- Otros impuestos menos subvenciones sobre la producción: 5
- Impuestos corrientes sobre la renta, el patrimonio, etc. pagados: 106

Sector Administraciones Públicas:

- Producción: 229
- Consumos Intermedios: 64
- Otros impuestos menos subvenciones sobre la producción (empleos) ≈ 0
- Impuestos recibidos:
 - Impuestos menos subvenciones sobre la producción y las importaciones: 130
 - Impuestos corrientes sobre la renta, el patrimonio, etc.: 128
- Consumo colectivo efectivo: 95
- Transferencias sociales en especie (TSE): 138

Sectores Sociedades no Financieras e Instituciones Financieras:

- Producción: 1.663
- Consumos Intermedios: 960
- Otros impuestos menos subvenciones sobre la producción: 7
- Impuestos corrientes sobre la renta, el patrimonio, etc. pagados: 23

Resto del Mundo:

- Remuneración de los asalariados no residentes por empleadores residentes: 1
- Impuestos menos subvenciones sobre la producción y las importaciones pagados por unidades residentes al resto del mundo: -3
- Impuestos corrientes sobre la renta, el patrimonio, etc. netos procedentes del resto del mundo ≈ 0
- Rentas de la propiedad netas procedentes del resto del mundo: -4
- Cotizaciones sociales netas procedentes del resto del mundo: -1
- Prestaciones sociales distintas de las TSE netas procedentes del resto del mundo: 4
- Otras transferencias corrientes netas procedentes del resto del mundo: -15



Se pide:

- (a) Calcular todos los saldos brutos de las cuentas corrientes del total de la economía para el año T .
- (b) Definir y calcular la tasa de ahorro nacional e indique algunos de los factores que influyen en ella.
- (c) Sabiendo que dicha tasa fue del 20,2 % en el año $T-5$, ¿qué conclusiones se podrían extraer a la luz del resultado obtenido?

Solución.

- (a) Para resolver este apartado trabajaremos sobre el cuadro siguiente, en el que los datos que directamente nos proporciona el problema relativos a los 5 sectores institucionales S.11, S.12, S.13, S.14 y S.15 están reflejados en gris, los datos naranjas son los que se refieren al sector Resto del Mundo y los datos en amarillo se calculan a partir de otros que son conocidos.

Por ejemplo, veamos dos datos en amarillo, empezando por la remuneración de asalariados que va en la cuenta de explotación, que es la interior, y que se calcula como

$$\begin{aligned} \text{RA}(\text{interior}) &= \text{RA}(\text{nacional}) - \text{pagada por no residentes a residentes} + \\ &\quad + \text{pagada por residentes a no residentes} = \\ &= 588(\text{dato del sector hogares}) - 3(\text{dato de hogares}) + \\ &\quad + 1(\text{dato de RM}) = 586 \end{aligned}$$

El dato D.2-D.3 que también está en amarillo, se obtiene del dato del sector AAPP que es 130 y al que hay que restarle el dato del resto del mundo -3 puesto que son impuestos que se pagan por unidades residentes al resto del mundo y por tanto en la cuenta de explotación no se consideran, puesto que en esta cuenta se trabaja con el concepto de interior.

Los datos en naranja proceden del resto del mundo. El dato de los impuestos corrientes sobre la renta, el patrimonio, etc, netos procedentes del resto del mundo, que es aproximadamente cero no aparece reflejado en la cuenta de distribución secundaria de la renta puesto que no afecta, pero si no fuera despreciable habría que considerarlo ahí.

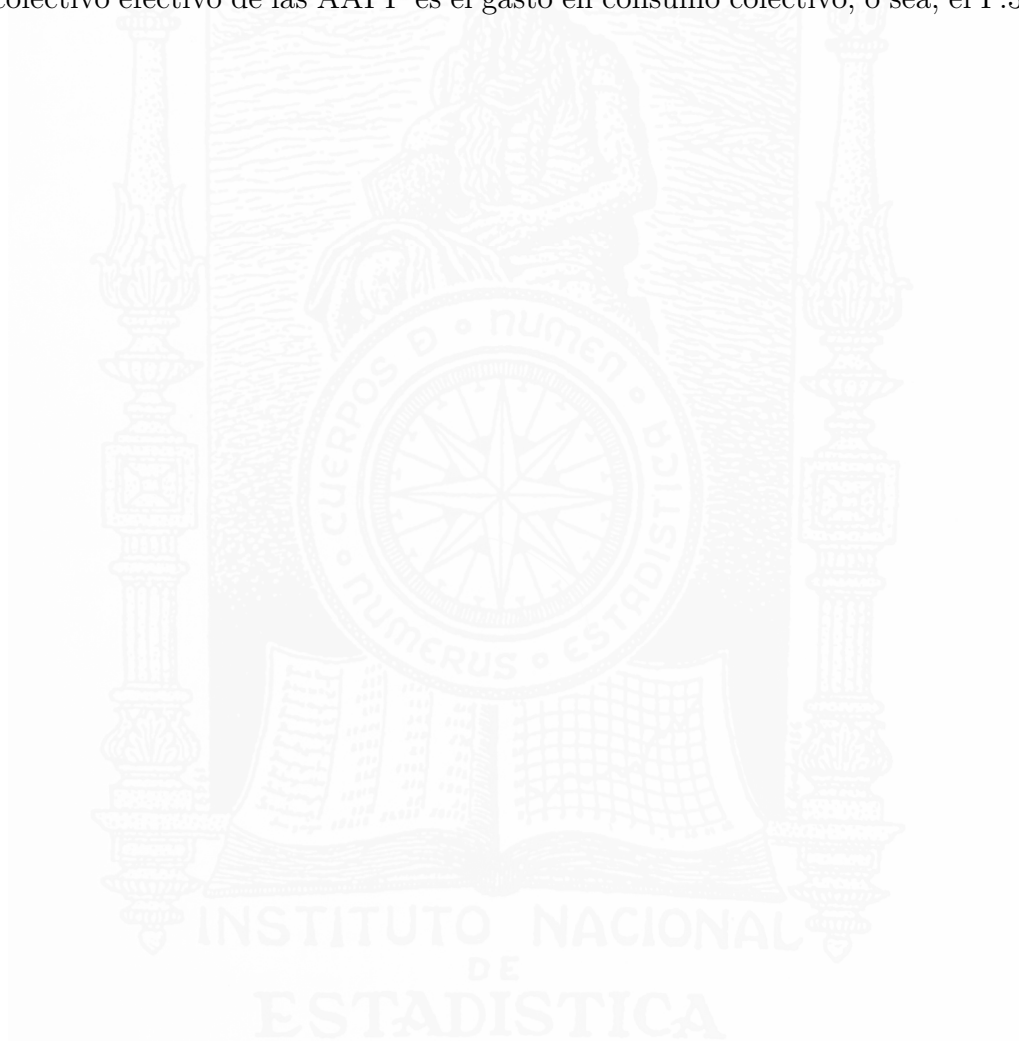
Hay un dato en la cuenta de producción, que es el de D.21-D.31 correspondiente a los impuestos menos subvenciones sobre los productos, que vale 115, puesto





que sabemos que $D.2-D.3 = 127$, celda amarilla, como hemos comentado, y $D.29-D.39 = 12$ se obtiene de los datos del problema directamente, por lo que por diferencia tenemos el valor de dichos impuestos netos.

En la cuenta de utilización de la renta disponible ponemos como dato de las administraciones públicas el valor 138, que son las transferencias sociales en especie, que es el gasto en consumo individual de este sector, es decir, el valor P.31, que coincide con el valor D.63 y es un dato que nos proporciona el problema. Todo el consumo de los hogares es individual, por lo que el dato del sector hogares de gasto en consumo final se refiere a individual. El consumo colectivo efectivo de las AAPP es el gasto en consumo colectivo, o sea, el P.32.





Capítulo 2.2. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte Común.

Cuentas del total de la economía y de los sectores institucionales Cuentas corrientes y cuentas de acumulación.

Unidad: millones de euros

Lista Tablas

Cuentas corrientes

I. Cuenta de producción

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
1.118	94	64	960	P.1	Producción	1.663	229	364	2.256
				P.2	Consumos intermedios				
				D.21-D.31	Impuestos menos subvenciones sobre los productos				115
1.253				B.1b/	Valor añadido bruto/				
				B.1'b	Producto interior bruto				

II. 1.1 Cuenta de explotación

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
586				B.1b/B.1'b	Valor añadido bruto / Producto interior bruto	703	165	270	1.253
127				D.1	Remuneración de los asalariados (Interior: 588-3+1)				
				D.2-D.3	Impuestos menos subvenciones sobre la producción y las importaciones (Interior)				
					(Nac = Int + RecibDelRM - PagAIRM, 130=Int + 0 - (-3))				
12	5	0	7	D.29-D.39	Otros impuestos menos subvenciones sobre la producción				
403				B.2b	Excedente de explotación bruto				
137	137			B.3b	Renta mixta bruta				

II. 1.2 Cuenta de asignación de la renta primaria

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
				B.2b	Excedente de explotación bruto				403
				B.3b	Renta mixta bruta				137
				D.1	Remuneración de los asalariados			588	588
				D.2 - D.3	Impuestos menos subvenciones sobre la producción y las importaciones (Nacional, que es el dato del problema)	130			130
				D.4	Rentas de la propiedad netas procedentes de RM				-4
1.254				B.5b/B.5'b	Saldo de rentas primarias bruto / Renta nacional bruta				

II. 2 Cuenta de distribución secundaria de la renta

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
129	106		23	B.5b/B.5'b	Saldo de rentas primarias bruto / Renta nacional bruta				1.254
				D.5	Impuestos corrientes sobre la renta, el patrimonio, etc.		128		128
				D.61	Cotizaciones sociales netas procedentes del RM				-1
				D.62	Prestaciones sociales netas procedentes del RM distintas de las transferencias sociales en especie				4
				D.7	Otras transferencias corrientes netas procedentes del RM				-15
1.241				B.6b	Renta disponible bruta				

II. 3 Cuenta de redistribución de la renta en especie

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
138		138		B.6b	Renta disponible bruta				1.241
				D.63	Transferencias sociales en especie			138	138
1.241				B.7b	Renta disponible ajustada bruta				

II. 4.1 Cuenta de utilización de la renta disponible

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
953				B.8b	Renta disponible bruta				1.241
858	720		138	P.3	Gasto en consumo final				
95			95	P.31	Gasto en consumo individual				
				P.32	Gasto en consumo colectivo				
				D.8	Ajuste por la variación de los derechos por pensiones (no afecta al total de la economía en la cuenta)				
288				B.8b	Ahorro bruto				

II. 4.2 Cuenta de utilización de la renta disponible ajustada

Empleos				Código	Operaciones y otros flujos y saldos contables	Recursos			
S.1	S.14-S.15	S.13	S.11-S.12			S.11-S.12	S.13	S.14-S.15	S.1
Total de la economía	Hogares e IPFLSH	Administraciones públicas	Sociedades no financieras- Instit. Financieras			Sociedades no financieras- Instit. Financieras	Administraciones públicas	Hogares e IPFLSH	Total de la economía
953				B.7b	Renta disponible ajustada bruta				1.241
858				P.4	Consumo final efectivo (Para el total de la economía es el GCF)				
95		95		P.41	Consumo individual efectivo				
				P.42	Consumo colectivo efectivo				
				D.8	Ajuste por la variación de los derechos por pensiones (no afecta al total de la economía en la cuenta)				
288				B.8b	Ahorro bruto				





- (b) En el ámbito de las cuentas nacionales, la tasa de ahorro nacional es un indicador macroeconómico que refleja la proporción del ingreso nacional disponible que no se destina al consumo, sino que se ahorra en una economía durante un período determinado (generalmente un año). La tasa de ahorro nacional muestra qué parte de los ingresos generados en un país se reserva para inversiones futuras, en lugar de ser consumida inmediatamente. Incluye el ahorro tanto del sector privado como del sector público. La fórmula de cálculo es:

$$\begin{aligned}\text{Tasa de ahorro nacional} &= \frac{\text{Ahorro nacional bruto}}{\text{Renta disponible bruta}} \cdot 100 = \frac{\text{B.8b}}{\text{B.6b}} = \\ &= \frac{288}{1241} \cdot 100 \approx 23,21 \%\end{aligned}$$

La tasa de ahorro nacional, calculada como el cociente entre el ahorro bruto (B.8b) y la renta disponible bruta (B.6b), está influida por diversos factores que actúan sobre los saldos de las cuentas corrientes del sistema SEC 2010. Uno de los principales determinantes es el nivel del gasto en consumo final (P.3), ya que un mayor consumo reduce directamente el ahorro bruto, dado que B.8b se define como B.6b menos P.3. Por tanto, una alta propensión al consumo tiende a disminuir la tasa de ahorro.

Otro factor relevante es el nivel de la renta disponible bruta (B.6b), que refleja la capacidad agregada de la economía para consumir y ahorrar. Esta renta se ve afectada por el ingreso primario bruto (B.5b), que incluye las remuneraciones, rentas de la propiedad y excedentes, así como por las transferencias corrientes netas recibidas del resto del mundo (D.7). En este sentido, las transferencias corrientes (como remesas o ayudas internacionales) pueden incrementar la renta disponible bruta y, en la medida en que no se consuman completamente, aumentar también el ahorro.

Las rentas de la propiedad netas, como parte del ingreso primario, son especialmente relevantes en economías con fuertes vínculos financieros internacionales. Una posición neta acreedora puede generar flujos de ingreso primario que aumenten la renta disponible. Asimismo, la política fiscal influye en la tasa de ahorro a través del impacto de las transferencias públicas y del consumo final del gobierno sobre los saldos agregados. Por ejemplo, un aumento del consumo público incrementa el gasto total (P.3), lo que reduce el ahorro, mientras que transferencias netas a los hogares pueden aumentar la renta disponible si no son compensadas por mayor consumo.



La estructura institucional y el comportamiento financiero de los sectores institucionales también desempeñan un papel importante. El ahorro nacional es la suma de los ahorros de los hogares, las sociedades no financieras, las sociedades financieras, las administraciones públicas y las ISFLSH. Por tanto, déficits públicos, bajas tasas de ahorro de los hogares o decisiones empresariales sobre distribución de beneficios influyen en el ahorro agregado de la economía.

Finalmente, existen factores estructurales y demográficos, como el envejecimiento poblacional o las preferencias culturales entre consumo y ahorro, que aunque no forman parte explícita de los saldos contables, afectan a su evolución a lo largo del tiempo. Asimismo, las relaciones económicas con el resto del mundo —en términos de ingresos primarios o transferencias— pueden modificar sustancialmente la renta disponible nacional y, por ende, la tasa de ahorro si los flujos netos recibidos no se traducen completamente en mayor consumo.

- (c) Un aumento de la tasa de ahorro nacional del 20,2 % en el año T-5 al 23,21 % en el año T representa un cambio significativo en el comportamiento macroeconómico de la economía. A la luz del marco contable del SEC 2010, este salto sugiere que han ocurrido transformaciones sustanciales en los saldos de las cuentas corrientes, particularmente en la renta disponible bruta (B.6b) y/o en el gasto en consumo final (P.3), que son los dos elementos que determinan directamente el ahorro bruto (B.8b).

Una primera hipótesis plausible es una reducción significativa del consumo total —de los hogares, de las ISFLSH o de las administraciones públicas— manteniéndose estable o con menor crecimiento la renta disponible. Esto podría deberse a periodos de incertidumbre económica (crisis, pandemia, shocks externos), donde los agentes tienden a ahorrar más y consumir menos por precaución. También podría estar asociado a ajustes fiscales que reducen el consumo público, especialmente si el gasto público representa una proporción elevada del total del consumo en la economía, o a cambios estructurales como el envejecimiento de la población o nuevas pautas culturales de comportamiento.

En segundo lugar, el aumento de la tasa de ahorro podría deberse a un crecimiento notable de la renta disponible bruta (B.6b). Este crecimiento puede deberse a incrementos del ingreso primario bruto (B.5b), por mejoras en salarios, rentas de la propiedad o beneficios empresariales, o bien a un aumento de transferencias netas recibidas del resto del mundo (D.7), como flujos de remesas, ayuda externa o rentas del capital provenientes del extranjero. También



podrían influir reformas estructurales o una mejora en la productividad que eleven la capacidad generadora de renta sin que ello se traduzca inmediatamente en un aumento del consumo.

Una tercera posibilidad es que haya ocurrido una contracción o control del consumo privado acompañado de una estabilidad o mejora de los ingresos. Un cambio voluntario en los patrones de consumo de los hogares, impulsado por razones culturales, políticas de austeridad, incertidumbre o expectativas desfavorables, puede hacer que una parte creciente de la renta se destine al ahorro.

Asimismo, una política fiscal contractiva o una consolidación presupuestaria puede haber influido significativamente. Una fuerte reducción del déficit público —por ejemplo, disminuyendo el consumo público y aumentando impuestos— puede mejorar el ahorro de las administraciones públicas, lo cual se traduce directamente en un mayor ahorro nacional si el sector privado no lo compensa con un aumento del consumo.

Por último, no puede descartarse la influencia de acontecimientos extraordinarios, como una pandemia, una crisis externa, un shock energético o un conflicto geopolítico. En este tipo de contextos, el consumo puede contraerse bruscamente por restricciones internas o externas, mientras que los ingresos pueden mantenerse estables gracias a transferencias públicas o externas, lo que genera un aumento transitorio pero pronunciado del ahorro.

En conclusión, el incremento de la tasa de ahorro nacional observado entre los años T-5 y T refleja una transformación económica relevante. Aunque no es posible atribuir el cambio a un único factor sin información adicional, es probable que responda a una combinación de menor consumo relativo, mayor renta disponible y factores extraordinarios o estructurales que hayan alterado el comportamiento agregado de los sectores institucionales. Para identificar con mayor precisión el origen del cambio, sería conveniente analizar la evolución de los valores absolutos de B.6b, P.3 y B.8b, así como la contribución de cada sector institucional al ahorro agregado.



2.3. Soluciones del cuarto examen del Cuerpo Superior de Estadísticos del Estado de 2024. Especialidad ciencia de datos. Turno libre.

Este examen se celebra el 15 de Septiembre de 2025, por lo que a la fecha de publicación de esta versión todavía no estaba disponible para poder resolverlo. La próxima versión de este ejemplar contendrá las soluciones de dicho examen.







Capítulo 3

Año 2024. Promoción interna.

"I dream my painting and I paint my dream"
—Vincent Willem Van Gogh.



Durante su internamiento voluntario en el asilo de Saint-Paul-de-Mausole en Saint-Rémy-de-Provence, el genio atormentado no abandonó su pasión por el arte, sino que la cultivó aún más. A pesar de las luchas mentales y emocionales que enfrentaba, Van Gogh encontró en la pintura una vía de escape y expresión. Durante este período, produjo algunas de sus obras más célebres y emblemáticas, incluyendo la cautivadora "Noche estrellada". A través de su arte en el asilo, Van Gogh demostró la fuerza resiliente de su espíritu creativo. En lugar de rendirse ante la adversidad, utilizó su pincel para traducir sus sentimientos y experiencias internas en obras de arte cautivadoras y conmovedoras. El arte de Van Gogh desde el manicomio trasciende las limitaciones físicas y mentales, dejando un legado duradero que nos recuerda la capacidad del arte para conectar con lo más profundo del ser humano.

$$2024 = 2^3 \cdot 11 \cdot 23$$

2024 is a Cunningham number, because is equal to $45^2 - 1$.



3.1. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024.

Parte común. Promoción interna.

Este examen tiene 6 ejercicios, de los cuales 5 de ellos aparecen en el examen de turno libre del año 2024 salvo el ejercicio número 6, o cuestión 6, que resolvemos a continuación.





Cuestión 6. El ayuntamiento de una ciudad tiene como objetivo estudiar el nivel de actividad de sus comercios locales y para ello ha diseñado una muestra probabilística. Cada comercio seleccionado ha respondido a un cuestionario y finalmente se dispone de los siguientes datos:

Variable	Tipo	Descripción
ID_Comercio	Alfanumérico	Código único de cada comercio. El identificador puede ser cualquier cadena de caracteres que asegure su unicidad, de 4 posiciones
Mes	Entero (1–12)	Mes de referencia de la actividad, donde 1 = Enero, 2 = Febrero, ..., 12 = Diciembre.
Actividad	Binaria (0,1)	Indica si el comercio ha tenido actividad durante el mes (1 = Sí, 0 = No).
Ventas.Tienda (€)	Numérica	Ventas realizadas en el local físico. El valor debe ser 0 o un número positivo.
Ventas.Online (€)	Numérica	Ventas realizadas a través de internet. El valor debe ser 0 o un número positivo.
Ventas.Totales (€)	Numérica	Total de ventas del mes, incluyendo las realizadas en tienda física y online . El valor debe ser 0 o un número positivo.
Gastos (€)	Numérica	Gastos del negocio durante el mes, incluyendo los gastos relacionados con los empleados (salarios, seguros sociales, etc.), alquiler, suministros, entre otros. El valor debe ser 0 o un número positivo.
Empleados	Entero	Número de empleados contratados. El valor debe ser 0 o un número entero positivo.

Se pide:

- (a) Establecer 5 controles de validación (*edits*) para los datos recogidos en la encuesta, cubriendo al menos uno de cada uno de los siguientes tipos: rango, balance y bivalente.

Se debe:

- Escribir la regla de validación como una condición lógica o en pseudocódigo, determinando si es de rango, balance y bivalente.
 - Proporcionar un ejemplo donde se incumpla la regla y explicar brevemente por qué sería un error.
- (b) En los datos recogidos en la encuesta, se ha observado la existencia de valores faltantes en algunas variables de ciertos cuestionarios. Suponiendo que la información recogida corresponde a cinco comercios (registros), se requiere imputar los valores faltantes (NA) utilizando el método indicado y los datos disponibles en la siguiente tabla.



ID	Mes	Actividad	Ventas.Tienda	Ventas.Online	Ventas.Totales	Gastos	Empleados
C001	1	1	1200	300	1500	1100	3
C002	1	1	1000	NA	NA	850	2
C003	1	0	0	0	0	0	0
C004	1	1	1200	400	1600	NA	4
C005	1	1	1000	400	1400	1300	5

Métodos de imputación a aplicar:

- **C002 - Ventas.Online:** imputación por la media.
- **C004 - Gastos:** vecino más próximo o cercano.

Indicar para cada caso el valor imputado y los pasos llevados a cabo para obtener dicho valor.

- (c) Suponga que se está valorando la posibilidad de difundir un fichero de microdatos que incluya las variables ventas totales , número de empleados e identificador del comercio para su explotación por parte de usuarios externos. Proponga al menos dos medidas concretas de control del secreto estadístico que podrían aplicarse sobre dichas variables antes de la difusión del fichero.

Solución.

- (a) Para resolver el apartado (a), se definen 5 controles de validación (edits) para los datos recogidos en la encuesta, abarcando al menos uno de cada uno de los siguientes tipos: rango, balance y bivariantes, cuya definición, de una forma resumida, es la siguiente:

- Edit de rango: verifica que un valor esté dentro de un intervalo válido.
- Edit de balance: verifica relaciones aritméticas entre variables (por ejemplo, suma o diferencia).
- Edit bivalente: verifica coherencia lógica entre dos variables (no necesariamente aritmética).

Vamos a definir dichos edits para los datos de la encuesta:

1. Edit de rango – Mes:

- Regla lógica: $1 \leq \text{Mes} \leq 12$
- Condición: el mes debe representar un valor válido del calendario.
- Ejemplo de error: si $\text{Mes} = 15$, es inválido porque no existe un mes número 15.



2. Edit de balance – Ventas.Totales (€):

- Regla lógica: $\text{Ventas.Totales} = \text{Ventas.Tienda} + \text{Ventas.Online}$
- Condición: las ventas totales deben ser la suma de ventas físicas y online.
- Ejemplo de error: si $\text{Ventas.Totales} = 2000$, $\text{Ventas.Tienda} = 1500$ y $\text{Ventas.Online} = 300$, la suma da $1800 \neq 2000$, lo cual es un error.

3. Edit bivalente – Actividad y Ventas:

- Regla lógica: si $\text{Actividad} = 0$, entonces $\text{Ventas.Tienda} = 0$ y $\text{Ventas.Online} = 0$
- Condición: si no hubo actividad en el mes, no deberían registrarse ventas.
- Ejemplo de error: si $\text{Actividad} = 0$ y $\text{Ventas.Online} = 100$, hay un error porque hubo ventas sin actividad.

4. Edit de rango – Empleados:

- Regla lógica: $\text{Empleados} \geq 0$ y debe ser un número entero.
- Condición: no puede haber empleados negativos y el número debe ser entero.
- Ejemplo de error: $\text{Empleados} = -2$ es inválido.

5. Edit bivalente – Gastos y Empleados:

- Regla lógica: si $\text{Empleados} > 0$, entonces $\text{Gastos} > 0$
- Condición: si hay empleados contratados, debe haber algún gasto (por salarios, seguros, etc.).
- Ejemplo de error: si $\text{Empleados} = 5$ y $\text{Gastos} = 0$, es incoherente.

(b) Vamos a detallar los pasos realizados para imputar cada uno de los siguientes valores, que son los que pide el enunciado según la tabla disponible:

- C002 - Ventas.Online: imputación por la media.
- C004 - Gastos: imputación por el vecino más próximo.

Analizamos cada caso por separado:

1. Caso C002 – Variable **Ventas.Online** (método: imputación por la media)

Se utiliza la media de la variable Ventas.Online de los comercios que tienen un valor observado distinto de NA y que tienen actividad, puesto





que el cuarto registro, aunque no tiene ningún valor vacío, están todos a cero puesto que la actividad es nula. Por tanto los valores disponibles son:

- C001: 300
- C004: 400
- C005: 400

Se calcula la media:

$$\text{Media} = \frac{300 + 400 + 400}{3} = \frac{1100}{3} \approx 366,67$$

Lo que hacemos es asignar el valor 366.67 a `Ventas.Online` de C002.

Además, se puede recalcular el valor de `Ventas.Totales` como:

$$\begin{aligned}\text{Ventas.Totales}_{C002} &= \text{Ventas.Tienda} + \text{Ventas.Online imputado} = \\ &= 1000 + 366,67 = 1366,67\end{aligned}$$

2. Caso C004 – Variable `Gastos` (método: vecino más próximo)

La utilización del método del vecino más próximo requiere de la definición de una medida de distancia entre registros utilizando las variables numéricas comunes: `Mes`, `Actividad`, `Ventas.Tienda`, `Ventas.Online`, `Empleados`. Sería necesario normalizar estas variables para evitar que una domine la distancia por su escala, pero por simplicidad realizaremos los cálculos sin normalizar.

Se calcula la distancia euclidiana entre C004 y los otros registros completos (sin NA en `Gastos`), teniendo en cuenta como comentábamos que trabajaremos sobre aquellos registros que presentan actividad y tienen gastos, que son:

- C001: `Gastos` = 1100
- C002: `Gastos` = 850
- C005: `Gastos` = 1300

Tomamos como ejemplo una distancia simplificada (sin normalizar, aunque idealmente se debe hacer), en la que utilizamos dos observaciones i y j y las tres variables disponibles que podemos incluir en la distancia, que son `Ventas.Tienda`(X_1), `Ventas.Online`(X_2) y `Empleados`(X_3) para definir la distancia entre ellas:



$$d(i, j) \stackrel{\text{not}}{=} d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{ik} - x_{jk})^2}$$

Para cada observación:

- C004 vs C001:

$$\begin{aligned} d_{41} &= \sqrt{(1200 - 1200)^2 + (400 - 300)^2 + (4 - 3)^2} = \\ &= \sqrt{0 + 100^2 + 1} = \sqrt{10000 + 1} \approx 100,005 \end{aligned}$$

- C004 vs C002:

$$\begin{aligned} d_{42} &= \sqrt{(1200 - 1000)^2 + (400 - 366,67)^2 + (4 - 2)^2} = \\ &= \sqrt{200^2 + 33,33^2 + 4} \approx \sqrt{41115} \approx 202,77 \end{aligned}$$

- C004 vs C005:

$$\begin{aligned} d_{45} &= \sqrt{(1200 - 1000)^2 + (400 - 400)^2 + (4 - 5)^2} = \\ &= \sqrt{200^2 + 0 + 1} = \sqrt{40000 + 1} \approx 200,002 \end{aligned}$$

El vecino más próximo es C001. Por tanto, se imputa el valor de Gastos = 1100 para C004.

- (c) Dado que el fichero de microdatos a difundir incluirá variables como **ventas totales**, **número de empleados** y el **identificador del comercio**, es necesario implementar medidas avanzadas de control del secreto estadístico. La inclusión del identificador hace imposible aplicar medidas clásicas de anonimización como la supresión directa, por lo que se deben aplicar técnicas que garanticen la confidencialidad sin comprometer la utilidad analítica del fichero. Este riesgo es especialmente relevante cuando los microdatos pueden ser cruzados con otras variables contextuales como la *rama de actividad*, la *comunidad autónoma* o la *provincia*, ya que existen sectores muy específicos donde unas pocas empresas concentran una gran parte de la actividad económica. En estos casos, una desanonimización mediante técnicas de cruce con bases de datos externas puede revelar información confidencial. A continuación, se proponen dos medidas concretas de control, aplicables de forma conjunta:

1. Top-coding y agrupamiento en variables cuantitativas

Esta técnica consiste en limitar la divulgación de valores extremos que podrían ser identificativos por su magnitud. Se aplica sobre variables como **ventas totales** y **número de empleados**, especialmente cuando:





- El valor se sitúa en el percentil 95 o superior del conjunto de datos.
- Existen combinaciones de actividad y localización que hacen que esos valores altos correspondan con alta probabilidad a una sola empresa.

Por tanto, se propone aplicar:

- *Top-coding en ventas totales*: los valores superiores al percentil 95 se reemplazan por dicho valor o se agrupan en un intervalo abierto (por ejemplo: “más de 250,000 €”).
- *Agrupamiento en número de empleados*: en lugar de informar el número exacto, se reemplaza por rangos predefinidos (por ejemplo: 0–2, 3–5, 6–10, ¿10).

Esto reduce la precisión para valores extremos, pero mantiene la capacidad de análisis agregada sin comprometer la confidencialidad de empresas singulares.

2. Evaluación de riesgo de reidentificación y perturbación dirigida

Dado que el fichero mantiene el identificador del comercio, se debe aplicar una evaluación formal del riesgo de reidentificación. Esto implica:

- Aplicar técnicas de *k-anonimato*: asegurar que cada combinación de variables quasi-identificadoras (como mes, actividad, empleados, ventas) aparezca al menos k veces en el conjunto de datos. Si no se cumple, se deben agrupar valores o aplicar perturbaciones.
- Perturbación dirigida sobre registros de alto riesgo: aquellos que no alcanzan el umbral k se modifican ligeramente mediante ruido controlado. Por ejemplo, se aplica un factor multiplicativo del tipo $X^* = X \cdot \epsilon$, con $\epsilon \in [0,95, 1,05]$, sobre **ventas totales**, de forma que los totales se mantengan estables en media, pero los valores individuales cambien suficientemente para evitar identificación.
- Asegurar que las distorsiones introducidas mantengan la coherencia contable del fichero (por ejemplo: no romper la relación **Ventas.Totales** = **Ventas.Tienda** + **Ventas.Online** si estas también se divulgan).

Este procedimiento se puede implementar con herramientas como *sdcMicro*, que permite evaluar el riesgo y simular ataques de linkage con bases auxiliares.

Finalmente, se recomienda acompañar la difusión del fichero con una documentación metodológica clara donde se describan las transformaciones aplicadas



Capítulo 3.1. Soluciones del segundo examen del Cuerpo Superior de Estadísticos del Estado de 2024. Parte común. Promoción interna.

y sus justificaciones técnicas, sin revelar detalles que permitan revertir los procesos de perturbación.





3.2. Soluciones del cuarto examen del Cuerpo Superior de Estadísticos del Estado de 2024. Especialidad ciencia de datos. Promoción interna.

Este examen, al igual que el de libre, a la fecha de publicación de esta versión del libro no se había celebrado, por lo que en los próximos meses se actualizará y aparecerán las soluciones de los ejercicios.

3.3. Soluciones del segundo examen del Cuerpo de Diplomados en Estadística del Estado de 2024. Promoción interna.

Todas las cuestiones de este examen aparecen en el que se ha puesto para el turno libre, por lo que las soluciones del mismo ya se han podido desarrollar en la sección [2](#).



Bibliografía

- Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. John Wiley.
- Ash, R.B. (1970). *Basic Probability Theory*. John Wiley.
- Aznar, G.A., García, A, y Martín, A.(1994). *Ejercicios de econometría I y II*. Pirámide
- Casella, G. y Berger, R.L. (1992). *Statistical Inference* (2nd ed.). Duxbury Advanced Series.
- Castro, J. (2009). *Statistical disclosure control in tabular data* (Report DR 2009-11), Universitat Politècnica de Catalunya.
- Celestino Rey, F. (2010). *Historia de los cuerpos especiales de estadística de la Administración General del Estado, 1860-2010*. Visión Libros.
- Cuadras, C.M. (2002). *Problemas de Probabilidad y Estadística*. PPU Barcelona.
- Cuadras, C.M. (2012). *Nuevos métodos de análisis multivariante*.
- García, A., Quesada, V. (1998). *Lecciones de Cálculo de Probabilidades*. Díaz de Santos.
- Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. Vol. 1. John Wiley.
- Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. Vol.2. John Wiley.
- Fernández, C.C., y Fuentes, F. (1995). *Curso de Estadística Descriptiva. Teoría y Práctica*. Madrid: Ariel.
- Fernández, F.R. y Mayor Gallego, J. A. (1995). *Ejercicios y prácticas de muestreo en poblaciones finitas*. EUB. Ediciones Universitarias de Barcelona.
- Gibbons, J. and S. Chakraborti (2011). *Nonparametric Statistical Inference*. Chapman and Hall/CRC; 5^a ed.
- Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Lenz, R., Nordholt, E. S., Seri, G. and De Woldf, P. P. (2010). *Handbook on statistical disclosure control*.



3. BIBLIOGRAFÍA

ESSNet SDC, Eurostat.

Ibarrola, P., Pardo, L., Quesada, V. (1997). Teoría de la Probabilidad. Síntesis.

Leguina, J. (1989). Fundamentos de Demografía.

Keppel, G. y Zedeck, S. (1989). Data Analysis for Research. Analysis of variance and multiple regression/ correlation approaches. New York: Freeman and Company.

Lequiller, F. y Derek Blades, (2014). Understanding National Accounts, OECD.

Livi Bacci, M. (1993). Introducción a la demografía, Ariel.

Martín-Guzmán, P. y Martín Pliego, F. (1985). Curso Básico de Estadística Económica. Madrid: AC.

Martín Pliego, F.J. (1990). Curso práctico de estadística económica.

Matilla, M; Pérez, P y Sanz, B. (2017). Econometría y Predicción. McGraw-Hill

Merediz, A. (2004). Historia de la estadística oficial como institución pública en España.

Montero, J., Pardo, L, Morales, D., Quesada, V. (1988). Ejercicios y Problemas de Cálculo de Probabilidades. Díaz de Santos.

Peña, D. (1987). Estadística. Modelos y métodos 1: Fundamentos. Madrid: Alianza Universidad.

Peña, D. (1987). Estadística. Modelos y métodos 2. Modelos Lineales y Series Temporales. Madrid: Alianza Universidad.

Peña, D. (2002). Analisis de datos multivariantes. McGraw-Hill Interamericana de España/UNED.

Pérez L. C. (2005). Muestreo estadístico: conceptos y problemás resueltos. Alhambra.

Pérez L. C. (2014). Modelos Econometricos Con Datos de Panel: Conceptos y Ejercicios Resueltos. Createspace.

Pfeffermann, D. y Rao, C.R. (2009). Handbook of statistics. Vol. 29A Sample surveys: design, methods and applications, North-Holland.

Pressat, Roland (1987). “Diccionario de la Demografía”. Oikos-Tau.

Pressat, Roland (1983). “El análisis demográfico”. Fondo de cultura económica.

Queouille, M. (1949). “Approximation tests of correlation in time series”. Journal of the Royal statistical Society, B, 11, 18-84.

Ruíz Maya, L. y Martín Pliego, F.J. (2004). Fundamentos de inferencia estadística Ed. Thompson-Paraninfo.

Särndal, C.-E, Swensson, B. y Wretman, J. (1992). Model assisted survey sampling. Springer.

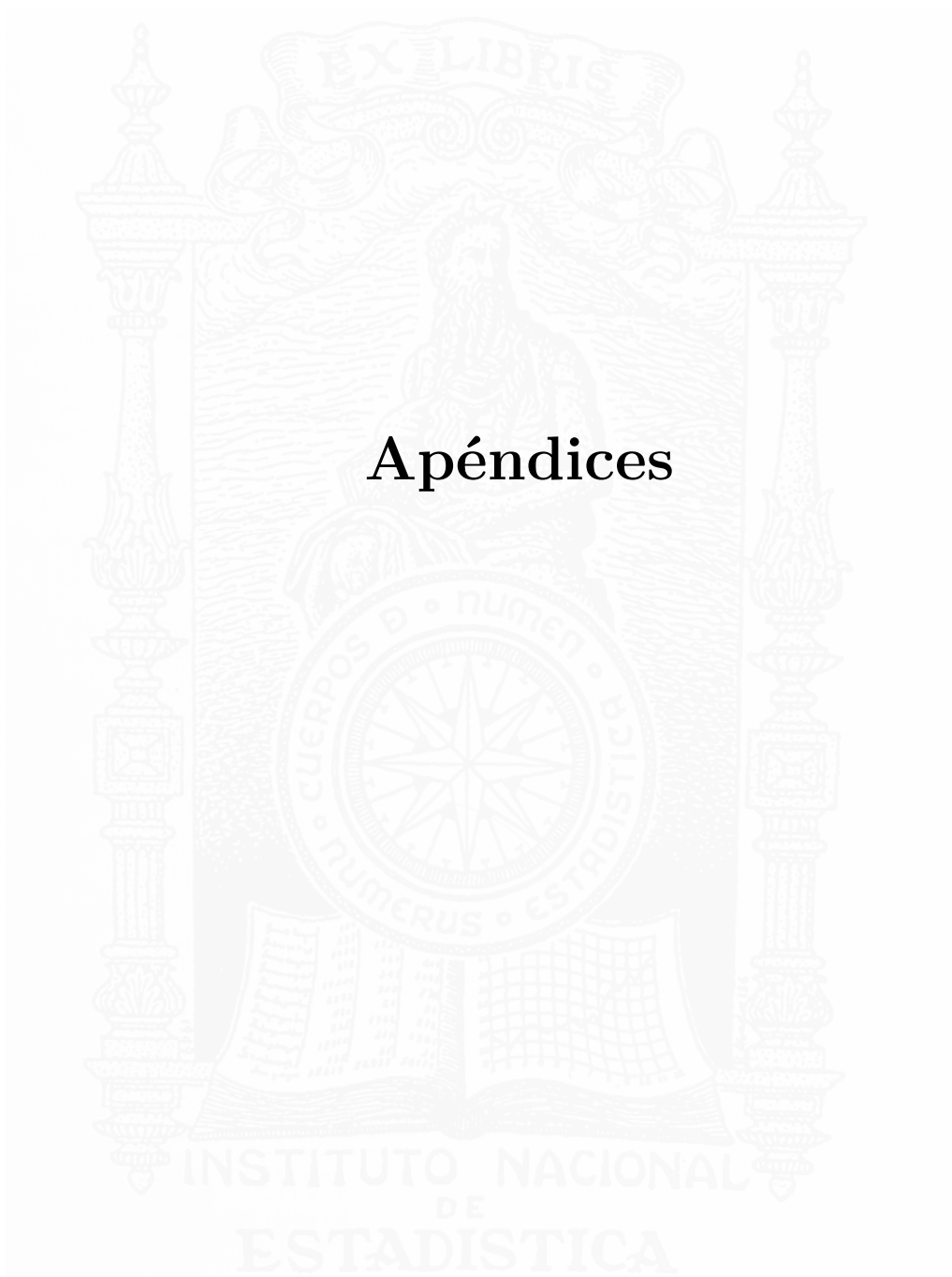


- Särndal, C.-E, y Lündstrom, S. (2005). Estimation in surveys with nonresponse. Wiley.
- Schattschneider, D., Escher, M. C. (2004). M.C. Escher: Visions of symmetry. New York: H.N. Abrams.
- Sistema Europeo de Cuentas, SEC 2010, (2013). EUROSTAT (Comisión Europea).
- Statistics Canada (2010). Survey methods and practices. Statistics Canada.
- Tukey, J. (1958). "Bias and confidence in not quite large samples". Annals of Mathematical Statistics, 29, 614.
- Uriel, E. y Peiró, A. (2000). Introducción al análisis de series temporales. Madrid: AC.
- Valliant, R., Dorfmann, A.H. y Royall, R.M. (2000). Finite population sampling and inference: a prediction approach. Wiley.
- Valliant, R. (1993). "Post-Stratification and Conditional Variance Estimation". Journal of the American Statistical Association, 88, 89-96.
- Vélez, R., Hernández, V. (1995). Cálculo de Probabilidades 1. UNED
- Vélez, R., y Garcia, A. (2012). Principios de Inferencia Estadística. UNED.
- Vinuesa, J. (1994). Demografía, análisis y proyecciones (Espacios y sociedades) .
- Vinuesa y Puga (2007). Técnicas y ejercicios de demografía.
- Wooldridge, J.B. (2001). Econometric Analysis of Cross-Section and Panel Data. MIT Press.
- Wooldridge, J.B. (2010). Introducción a la econometría. Un enfoque moderno. Cengage Learning.





Apéndices







Apéndice A

Exámenes de problemas de 2024. Turno libre.

- A.1. Segundo examen del Cuerpo Superior de Estadísticos del Estado. Parte común. Turno libre.



MINISTERIO
DE ECONOMÍA, COMERCIO
Y EMPRESA

INē

Instituto Nacional de Estadística

Oposición al Cuerpo Superior de Estadísticos de Estado

Segundo Ejercicio

Convocatoria de la oferta pública de empleo de 2024

Resolución de 22 de diciembre de 2024, de la Subsecretaría, por la que se convocan procesos selectivos para ingreso, por el sistema de acceso libre y promoción interna, en el Cuerpo Superior de Estadísticos del Estado. (BOE 31 de Diciembre de 2024)



Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 1. Se considera una población finita de tamaño N , compuesta por unidades ordenadas de acuerdo con una variable de interés definida por:

$$y_k = k, \quad \text{para } k = 1, 2, \dots, N.$$

donde los parámetros poblacionales de esta variable son:

- Media: $\mu_y = \frac{N+1}{2}$,
- Varianza: $\sigma_y^2 = \frac{(N-1)(N+1)}{12}$.

Además, se supone que $N = n \cdot a$, con $n, a \in \mathbb{N}$.

El objetivo es analizar la varianza del estimador de Horvitz-Thompson de la media poblacional bajo distintos esquemas de muestreo. Para ello, se plantea:

- (a) Suponiendo que se selecciona una muestra aleatoria simple sin reemplazamiento de tamaño n , deduzca la expresión de la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (b) La población se divide en n estratos consecutivos de tamaño a , de acuerdo al orden según la variable de interés, extrayéndose una unidad aleatoriamente por estrato. Deduzca la varianza del estimador de Horvitz-Thompson de la media poblacional en el muestreo estratificado descrito, sabiendo que la varianza poblacional dentro de cada estrato es $\sigma_i^2 = \frac{a^2-1}{12}, \forall i = 1, \dots, n$.
- (c) Suponer que se selecciona una muestra de tamaño n mediante muestreo sistemático con probabilidades iguales aplicado sobre la población ordenada de acuerdo a la variable de interés. Obtenga la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (d) Compare los tres planes de muestreo y justifique cuál resulta más adecuado en la situación planteada.

Observación:

Sea $n \in \mathbb{N}$. Entonces se verifican las siguientes identidades:

- La suma de los primeros n números naturales está dada por:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

- La suma de los cuadrados de los primeros n números naturales está dada por :

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$





Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 2. Un investigador está analizando la duración (en horas) de cierto tipo de bombillas LED. Se asume que la duración de las bombillas sigue una distribución exponencial con parámetro λ . Se toma una muestra aleatoria simple de tamaño $n = 10$, con los siguientes datos (en horas):

$\{88, 102, 96, 91, 110, 85, 99, 107, 95, 100\}$

El fabricante afirma que el tiempo medio de vida es de al menos 100 horas.

- (a) Obtenga en el contexto de este problema, el valor del parámetro λ y qué significado tiene el mismo.
- (b) Formule el contraste de hipótesis para comprobar si los datos contradicen la afirmación del fabricante utilizando un contraste de razón de verosimilitudes.
- (c) Calcule la función de verosimilitud bajo la hipótesis nula y bajo la alternativa (en este caso usa el estimador máximo verosímil)
- (d) Calcule el estadístico de razón de verosimilitudes y determine si se rechaza H_0 al nivel $\alpha = 0,05$.
- (e) Interprete el resultado en el contexto del problema.



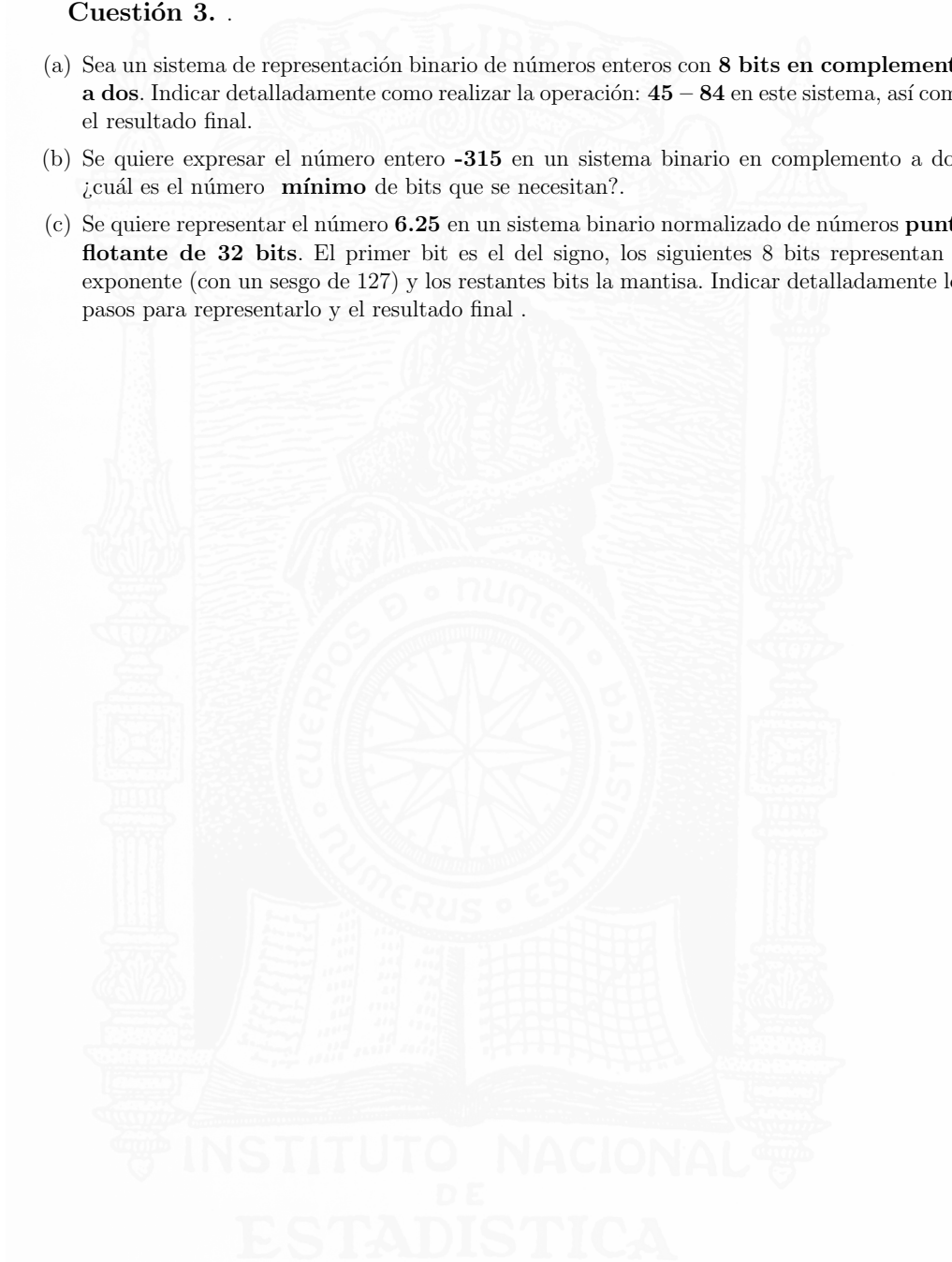
Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 3. .

- (a) Sea un sistema de representación binario de números enteros con **8 bits en complemento a dos**. Indicar detalladamente como realizar la operación: $45 - 84$ en este sistema, así como el resultado final.
- (b) Se quiere expresar el número entero **-315** en un sistema binario en complemento a dos, ¿cuál es el número **mínimo** de bits que se necesitan?.
- (c) Se quiere representar el número **6.25** en un sistema binario normalizado de números **punto flotante de 32 bits**. El primer bit es el del signo, los siguientes 8 bits representan el exponente (con un sesgo de 127) y los restantes bits la mantisa. Indicar detalladamente los pasos para representarlo y el resultado final .





Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 4. En un país P , se han registrado los datos de población por edad y de nacimientos por edad de la madre, que se proporcionan en la tabla. Las cifras están expresadas en miles de personas. A partir de esos datos, calcule los siguientes indicadores:

- (a) Tasa bruta de natalidad.
- (b) Ratio de masculinidad al nacimiento.
- (c) Tasa global de fecundidad.
- (d) Tasa específica de fecundidad para el grupo de edad “De 30 a 34 años”.
- (e) Índice sintético de fecundidad: indique dos formas de obtenerlo y calcúlelo utilizando la forma abreviada.
- (f) Tasa bruta de reproducción: indique dos formas de obtenerla y calcúlela utilizando la forma abreviada.

¿Qué información necesitaría para poder calcular la tasa neta de reproducción?

(Miles de personas)	Nacimientos 2023			Población 01/01/2023			Población 01/01/2024		
	Total	Hombres	Mujeres	Total	Hombres	Mujeres	Total	Hombres	Mujeres
Menos de 15	0	0	0	3.269	1.682	1.587	3.216	1.655	1.561
De 15 a 19	3	1	1	1.281	661	621	1.323	683	640
De 20 a 24	13	6	6	1.252	645	607	1.302	670	632
De 25 a 29	28	14	14	1.287	655	633	1.312	669	643
De 30 a 34	53	27	26	1.400	705	695	1.420	717	703
De 35 a 39	47	24	23	1.543	769	774	1.535	766	769
De 40 a 45	16	8	8	1.870	938	932	1.819	910	909
De 45 a 49	2	1	1	2.028	1.024	1.004	2.050	1.033	1.017
50 o más	0	0	0	10.115	4.706	5.409	10.337	4.813	5.524
De 15 a 49	160	82	78	10.659	5.395	5.264	10.758	5.447	5.312
Total	160	82	78	24.042	11.783	12.260	24.310	11.914	12.396



Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 5. La estimación más reciente del Producto Interior Bruto a precios de mercado (PIBpm) para el año 2023 de la Contabilidad Nacional Anual de España fue de 1.498 miles de millones €(mm€). Se dispone, además, de la siguiente información por ramas de actividad en las mismas unidades:

Agricultura, ganadería, silvicultura y pesca:	
Consumos Intermedios	38
Valor añadido bruto	37
Remuneración de los asalariados:	8
Otros impuestos netos sobre la producción:	-6
Industria:	
Consumos Intermedios	603
Remuneración de los asalariados:	98
Otros impuestos netos sobre la producción:	1
Construcción:	
Consumos Intermedios	122
Valor añadido bruto	81
Remuneración de los asalariados:	48
Otros impuestos netos sobre la producción:	3
Servicios:	
Consumos Intermedios	679
Valor añadido bruto	1.029
Remuneración de los asalariados:	562
Otros impuestos netos sobre la producción:	16
Otra información contable disponible:	
Formación bruta de capital:	315
Importaciones de bienes y servicios:	511
Saldo de intercambios exteriores de bienes y servicios	-59
Excedente de explotación bruto/Renta mixta bruta de la economía	639

Se pide:

- Calcular el valor añadido bruto de la industria.
- Elaborar la cuenta de bienes y servicios de la economía simplificada para ese año.
- Sabiendo que en 2023 la tasa de variación interanual del índice de volumen encadenado del PIBpm fue del 2,7 % y que el empleo equivalente a tiempo completo experimentó un aumento del 3,27 %. ¿Qué se podría decir del comportamiento de la productividad en ese año?

5





Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 6. Las autoridades municipales desean conocer la utilización de un nuevo tramo del carril bici.

Para ello, se empleó un muestreo por conglomerados sin reemplazamiento y con probabilidades iguales. En concreto, se seleccionaron aleatoriamente 4 semanas del total de 52 que componen el año. Durante cada una de las semanas seleccionadas, se registró diariamente el número de bicicletas que circularon por un punto específico del nuevo tramo del carril bici.

Los datos recogidos se presentan en la siguiente tabla:

Semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Total Semana
1	20	18	15	17	22	25	30	147
2	12	10	13	11	14	15	17	92
3	14	16	12	13	15	20	21	111
4	15	13	10	12	14	17	18	99

Se pide:

- Una estimación insesgada del total de usuarios anuales del nuevo tramo del carril bici y su error de muestreo relativo o coeficiente de variación. Calcule también un intervalo de confianza del 95 % para dicho total. Considerar $z_{\frac{\alpha}{2}} = 2$.
- Dichas autoridades estarían interesadas en estimar el número medio diario de usuarios del carril bici en un año futuro utilizando el mismo diseño muestral. Con el objetivo de garantizar que la varianza del estimador del número medio de usuarios diarios sea menor que uno, se pide determinar el número mínimo de semanas que deben seleccionarse.



Cuestión 7. Sean los siguientes modelos

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u \quad (\text{Modelo 1})$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u \quad (\text{Modelo 2})$$

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + u \quad (\text{Modelo 3})$$

donde las variables y, x_1, x_2, u son aleatorias tales que los modelos están correctamente definidos.

Se pide:

- (a) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, y obtenga el efecto parcial de x_2 sobre $E(y|x_1, x_2)$ para los modelos 1 y 2.
- (b) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, e interprete el significado de β_1 y β_2 en el modelo 3 (justifique la respuesta).
- (c) Obtenga, para el Modelo 1, $\mathbb{E}(u)$ sabiendo que $\mathbb{E}(u|x_1, x_2) = k$, $k \in \mathbb{R}$
- (d) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, obtenga (si es posible con la información facilitada) y para el Modelo 1 las siguientes covarianzas: $\text{cov}(u, x_1)$, $\text{cov}(u, x_2^2)$.
- (e) Explique el significado y alcance de $\mathbb{E}(u|x_1, x_2) = 0$ en el contexto de la modelización.





Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuestión 8. De las cuentas no financieras de los Sectores Institucionales de una Economía se conocen los siguientes datos correspondientes a un año T, en unidades monetarias:

Sector Hogares e ISFLSH:

- Producción: 364
- Consumos Intermedios: 94
- Gasto en consumo final: 720.
- Remuneración de los asalariados residentes: 588, de los que 3 son de empleadores no residentes.
- Renta mixta bruta: 137
- Otros impuestos menos subvenciones sobre la producción: 5
- Impuestos corrientes sobre la renta, el patrimonio, etc. pagados: 106

Sector Administraciones Públicas:

- Producción: 229
- Consumos Intermedios: 64
- Otros impuestos menos subvenciones sobre la producción (empleos) ≈ 0
- Impuestos recibidos:
 - Impuestos menos subvenciones sobre la producción y las importaciones: 130
 - Impuestos corrientes sobre la renta, el patrimonio, etc.: 128
- Consumo colectivo efectivo: 95
- Transferencias sociales en especie (TSE): 138

Sectores Sociedades no Financieras e Instituciones Financieras:

- Producción: 1.663
- Consumos Intermedios: 960
- Otros impuestos menos subvenciones sobre la producción: 7
- Impuestos corrientes sobre la renta, el patrimonio, etc. pagados: 23

Resto del Mundo:

- Remuneración de los asalariados no residentes por empleadores residentes: 1
- Impuestos menos subvenciones sobre la producción y las importaciones pagados por unidades residentes al resto del mundo: -3
- Impuestos corrientes sobre la renta, el patrimonio, etc. netos procedentes del resto del mundo ≈ 0
- Rentas de la propiedad netas procedentes del resto del mundo: -4
- Cotizaciones sociales netas procedentes del resto del mundo: -1
- Prestaciones sociales distintas de las TSE netas procedentes del resto del mundo: 4
- Otras transferencias corrientes netas procedentes del resto del mundo: -15



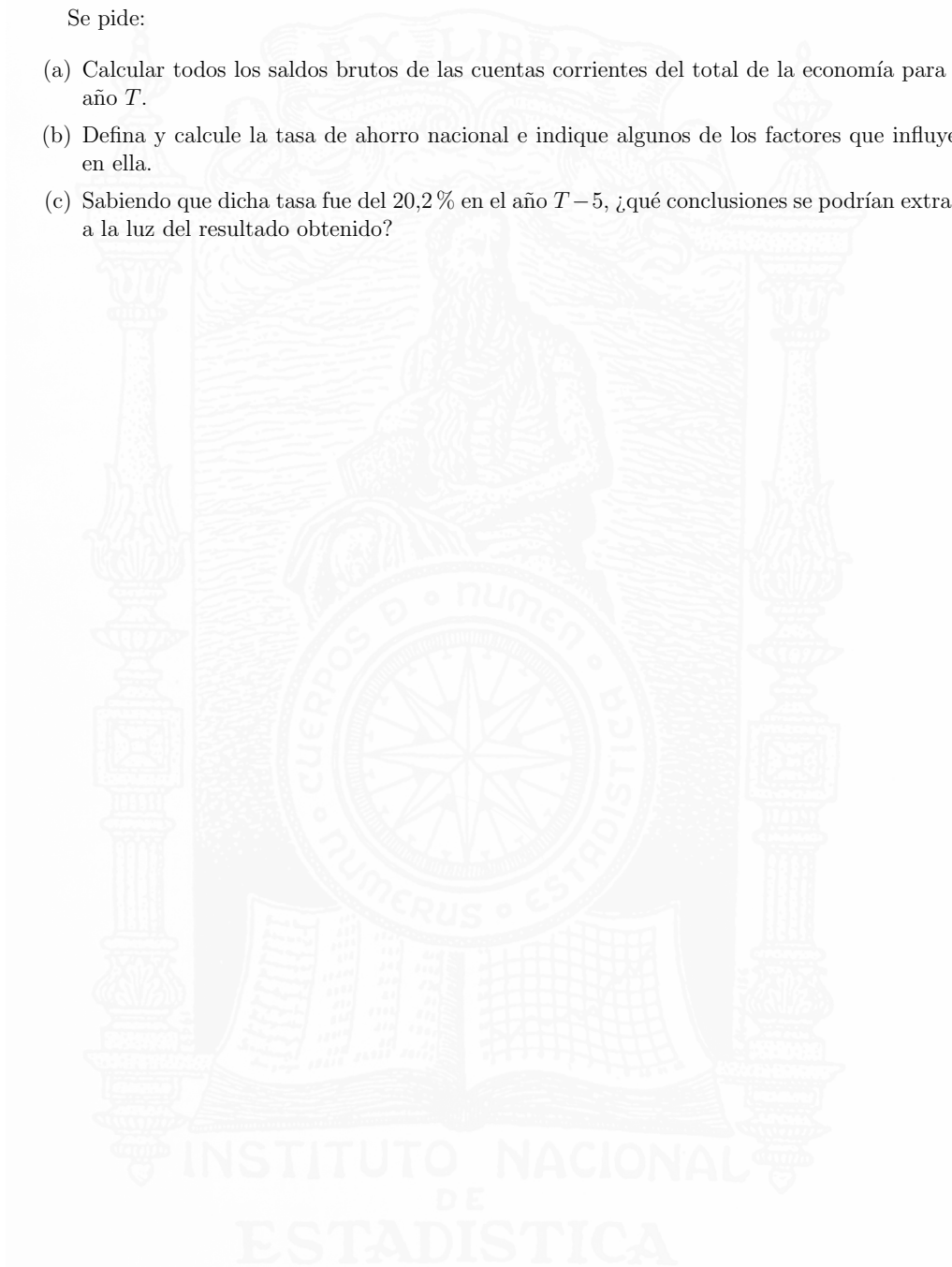
Apéndice A.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Se pide:

- (a) Calcular todos los saldos brutos de las cuentas corrientes del total de la economía para el año T .
- (b) Defina y calcule la tasa de ahorro nacional e indique algunos de los factores que influyen en ella.
- (c) Sabiendo que dicha tasa fue del 20,2 % en el año $T - 5$, ¿qué conclusiones se podrían extraer a la luz del resultado obtenido?





Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

A.2. Cuarto examen del Cuerpo Superior de Estadísticos del Estado. Especialidad ciencia de datos. Turno libre.

A.3. Segundo examen del Cuerpo de Diplomados en Estadística del Estado. Turno libre.





MINISTERIO
DE ECONOMÍA, COMERCIO
Y EMPRESA



**PROCESO SELECTIVO CONVOCADO POR RESOLUCIÓN DE 22 DE
DICIEMBRE DE 2024, DE LA SUBSECRETARÍA DE ECONOMÍA, COMERCIO Y
EMPRESA (BOE 30 DE DICIEMBRE DE 2024)**

**SEGUNDO EJERCICIO DE LA FASE DE OPOSICIÓN DE DIPLOMADOS EN ESTADÍSTICA
DEL ESTADO. SISTEMA GENERAL DE ACCESO LIBRE
CUESTIONARIO 2025/DEE-L**

AVISO IMPORTANTE PARA LOS OPOSITORES

1. De conformidad con lo establecido en las bases de la convocatoria, deberá elegir y contestar un máximo de **SIETE (7) cuestiones prácticas de las 10 planteadas**. En caso de que exceda el número máximo de cuestiones elegidas, el Tribunal solo corregirá las siete primeras preguntas respondidas, según el orden establecido en el cuestionario del examen, desechando en su caso, el resto de las respuestas. En caso de responder una pregunta que NO quiere que se le corrija, táchela claramente para que el Tribunal lo tenga en cuenta.
2. El tiempo máximo para la realización de este ejercicio es de **150 minutos**.





Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

CUESTIONES PRACTICAS BASADAS EN EL PROGRAMA DEL ANEXO I

Cuestión práctica 1.

Durante un apagón localizado en cuatro barrios de una ciudad española se han medido dos variables: la duración del apagón (en horas) y el número de hogares afectados por barrio, tal y como se puede ver en la siguiente tabla:

Duración del apagón (en horas)	Hogares del Barrio A	Hogares del Barrio B	Hogares del Barrio C	Hogares del Barrio D
1	10	15	12	13
2	20	30	24	26
3	30	45	36	39
4	40	60	48	52
5	50	75	60	65

Teniendo en cuenta estos datos:

- 1) Calcule las frecuencias marginales de los hogares afectados por el apagón en cada barrio.
- 2) Si nos fijamos en el barrio B, ¿Cuál es la frecuencia condicionada a que el apagón dure 5 horas?
- 3) ¿Cuál es la distribución de la duración del apagón si sabemos que los hogares afectados están en el barrio A?
- 4) Calcule la duración media del apagón en horas por hogar.

Cuestión práctica 2.

El consumo de leche de una guardería y los índices de precios de la leche en un determinado país en cinco años consecutivos han sido los siguientes:

Año	Consumo en litros de leche	Índice de precios de la Leche (Base 2021 = 100)
2020	900	90
2021	1.000	100
2022	1.100	102
2023	1.200	114
2024	1.100	120

El precio de la leche en el año 2024 asciende a 2 Euros/litro.

Calcule, a precios corrientes de cada año, el gasto en leche que ha tenido dicha guardería en cada uno de los cinco años mostrados.



Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

Cuestión práctica 3.

Sean X e Y dos variables aleatorias cuya función conjunta de densidad es la siguiente:

$$f_{XY}(x, y) = \begin{cases} k \cdot x, & 0 < x < y < 2 \\ 0, & \text{resto de casos} \end{cases}$$

Donde k es una constante de valor desconocido. Obtenga razonadamente:

- El valor de k .
- La distribución marginal de Y.
- $P(X < 0,5 / Y = 1)$
- Una aproximación de $P(\sum_{i=1}^{100} W_i < 100)$, siendo $W_1, W_2, \dots, W_{100}$, un total de 100 variables aleatorias independientes con idéntica distribución a X.

Cuestión práctica 4.

En un estudio estadístico se dispone de información de dos variables: X e Y. Se han resumido los datos en la siguiente tabla:

x_i	y_j	n_{ij}
10-30	0-10	3
10-30	10-20	3
30-50	10-20	2
50-110	20-30	2

Además, se han realizado algunos cálculos relativos a los momentos bidimensionales respecto al origen a_{rs} en relación a la citada información (r es el orden para la variable X, y s es el orden para la variable Y):

$$a_{10} = 36 \quad a_{20} = 1.840 \quad a_{01} = 14 \quad a_{02} = 245 \quad a_{11} = 640$$

Se está interesado en realizar predicciones de la variable Y utilizando la información de la variable X. Responda razonadamente a las siguientes preguntas:

- Obtenga la recta de regresión de Y sobre X que minimice la varianza residual.
- Considere las variables Z y W, que se definen como sigue:
 $Z=2X$ $W=Y+3$

Calcule la recta de regresión de W sobre Z que minimice la varianza residual.





Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

Cuestión práctica 5.

Un grupo de científicos militares trabajan desarrollando una bomba antes de que lo hagan sus enemigos. Los datos recolectados en 25 pruebas experimentales sugieren que la energía liberada (en kilotones) sigue una distribución normal de media y varianza desconocida.

Las mediciones obtenidas para la media y cuasidesviación típica muestral son las siguientes:

$$\bar{x} = 20,75 \text{ kilotones}$$

$$s_c = 1,25 \text{ kilotones}$$

- Calcule un intervalo de confianza del **95%** para la media de la energía liberada en las explosiones experimentales.
- Se necesita que la energía media liberada supere los **20 kilotones**. Realice un contraste de hipótesis con un nivel de significación del **5%**. ¿Deben los científicos continuar perfeccionando de la bomba, o la energía liberada es suficiente?
- Calcule la región de rechazo para la media, $\bar{x}$, con un nivel de significación de $\alpha = 0,01$. Tenga en cuenta que ahora se sabe que la desviación típica poblacional es $\sigma = 1,5$. La región de rechazo tiene la forma:
$$RC = \{Z > Z_{\alpha}\}$$
- Si exigimos que el valor superior de la función de potencia sea de **0,0013**, ¿qué media $\bar{x}$ de energía liberada necesitamos alcanzar en **4 pruebas extra** que se vuelven a realizar?

NOTA: Se proporcionan los siguientes valores que pueden ser de utilizad para la resolución de esta cuestión práctica.

Valores tabla t de Student

$$t_{24;0,025} = 2,064; t_{24;0,05} = 1,711; t_{25;0,025} = 2,060; t_{25;0,05} = 1,708$$

Valores tabla Normal

$$Z_{0,05} = 1,64; Z_{0,025} = 1,96; Z_{0,01} = 2,33; Z_{0,0013} = 3$$



Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

Cuestión práctica 6.

Se dispone de la siguiente información de una economía (en unidades monetarias):

	TOTAL u. m.
Producción	1.700
Consumos intermedios	700
Remuneración de los asalariados	525
Remuneración de los asalariados residentes que trabajan para unidades no residentes	17
Remuneración de los asalariados no residentes que trabajan para unidades residentes	7
Excedente Bruto de Explotación / Renta mixta	425
Impuestos	
- Impuestos sobre los productos	400
- Impuestos sobre la producción y las importaciones	350
Subvenciones	
- Otras subvenciones a los productos	300
- Subvenciones a la producción y las importaciones	200
Rentas netas de la propiedad pagadas al resto del mundo	15
Consumo final efectivo individual	450
Consumo final efectivo colectivo	350

Se pide calcular:

- Valor Añadido Bruto a precios básicos.
- Producto Interior Bruto a precios de mercado de la economía. Debe obtenerse utilizando el Valor Añadido Bruto del apartado anterior.
- Renta Nacional Bruta.
- Gasto en Consumo Final.
- Completar la Cuenta de producción y la Cuenta de explotación de la economía, incluyendo sus saldos.





Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

Cuestión práctica 7.

Considere un monopolista con costes totales $C(q) = 2q + 1$ que se enfrenta a una curva de demanda $P(q) = 10 - q$.

- Halle el equilibrio (precio y cantidad) del monopolio.
- Halle el excedente total del monopolio y la pérdida irrecuperable de eficiencia.
- Halle la función de costes medios e indique los valores de q para que este monopolio sea considerado un monopolio natural.
- La autoridad gubernamental propone una regla de fijación de precios consistente en igualar el precio al coste marginal. Cuantifique cómo varía el excedente total del mercado como consecuencia de dicha política.

Cuestión práctica 8.

En la oficina estadística de un país europeo se dispone de los siguientes datos:

	HOMBRES					MUJERES				
	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años
Población media residente en 2024 en miles de personas	9.500	1.900	4.750	1.900	950	10.000	1.800	4.300	2.300	1.600
Defunciones de residentes registradas durante el año 2024	78.755	380	2.375	9.500	66.500	88.370	180	1.290	6.900	80.000

Teniendo en cuenta estos datos, se pide:

- Calcule las tasas específicas de mortalidad por grupos de edad y sexo.
- Usando las tasas del apartado anterior, calcule la tasa bruta de mortalidad por sexo.
- Si tenemos en cuenta las tasas específicas de mortalidad del primer apartado y que la tasa bruta de mortalidad en mujeres es superior a la tasa bruta de mortalidad en hombres, explique cómo se comporta la mortalidad entre ambos sexos. ¿Por qué la mortalidad diferencial por sexo no tiene la misma tendencia en las tasas brutas por sexo que en las tasas específicas por sexo y edad? Explique una mejor forma de hacer la comparativa de las tasas por sexo en este caso.

NOTA: Puede dejar los cálculos de las tasas sin llegar a sus cifras decimales siempre que se indique de forma clara y concisa y cómo ha llegado hasta ellos.



Apéndice A.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado

Cuestión práctica 9.

Consideramos los sistemas de representación de números enteros magnitud y signo, complemento a uno y complemento a dos, con 8 bits.

- Explique las diferencias entre estos tres métodos.
- Represente el número entero **-93** con los tres métodos mencionados.
- En un sistema binario (sin bit para signo), ¿A qué valor decimal exacto equivale el número **101010,101**?

Cuestión práctica 10.

El propietario de una finca de naranjos está preocupado por los daños causados por una DANA, y para evaluar dicho desastre decide contratar los servicios de dos estadísticos, A y B, que realizarán un muestreo.

El dueño de la finca les proporciona la siguiente tabla, estratificada según la variedad:

Estratos (variedad)	Número de árboles de la población (N_h)	Proporción estimada de árboles afectados
1	10.000	0,7
2	6.000	0,4
3	4.000	0,15

Si el estadístico A es partidario del muestreo estratificado sin reemplazamiento con afijación proporcional, calcule:

- El tamaño de muestra total que debería utilizar para tener la misma precisión $e = 0,01$ con la que estimaría P en una muestra aleatoria simple sin estratificar.
- El tamaño muestral de cada estrato.

Por otra parte, el estadístico B prefiere el muestreo estratificado con reemplazamiento y afijación óptima. Sin embargo, su material de trabajo lo alquila.

Calcula que, para visitar cada estrato, los costes serán:

$$C_1 = 4 \text{ € por árbol}$$

$$C_2 = 4 \text{ € por árbol}$$

$$C_3 = 9 \text{ € por árbol}$$

- Calcular el tamaño muestral de cada estrato, si el tamaño de muestra total es el mismo que quiere utilizar el estadístico A.







Apéndice B

Exámenes de problemas de 2024. Promoción interna.

- B.1. Segundo examen del Cuerpo Superior de Estadísticos del Estado. Parte común. Turno de promoción interna.



Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.



MINISTERIO
DE ECONOMÍA, COMERCIO
Y EMPRESA

INē
Instituto Nacional de Estadística

Oposición al Cuerpo Superior de Estadísticos de Estado

Segundo Ejercicio

Convocatoria de la oferta pública de empleo de 2024

Resolución de 22 de diciembre de 2024, de la Subsecretaría, por la que se convocan procesos selectivos para ingreso, por el sistema de acceso libre y promoción interna, en el Cuerpo Superior de Estadísticos del Estado. (BOE 31 de Diciembre de 2024)

Acceso Promoción Interna, CDEE



Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 1. Se considera una población finita de tamaño N , compuesta por unidades ordenadas de acuerdo con una variable de interés definida por:

$$y_k = k, \quad \text{para } k = 1, 2, \dots, N.$$

donde los parámetros poblacionales de esta variable son:

- Media: $\mu_y = \frac{N+1}{2}$,
- Varianza: $\sigma_y^2 = \frac{(N-1)(N+1)}{12}$.

Además, se supone que $N = n \cdot a$, con $n, a \in \mathbb{N}$.

El objetivo es analizar la varianza del estimador de Horvitz-Thompson de la media poblacional bajo distintos esquemas de muestreo. Para ello, se plantea:

- (a) Suponiendo que se selecciona una muestra aleatoria simple sin reemplazamiento de tamaño n , deduzca la expresión de la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (b) La población se divide en n estratos consecutivos de tamaño a , de acuerdo al orden según la variable de interés, extrayéndose una unidad aleatoriamente por estrato. Deduzca la varianza del estimador de Horvitz-Thompson de la media poblacional en el muestreo estratificado descrito, sabiendo que la varianza poblacional dentro de cada estrato es $\sigma_i^2 = \frac{a^2-1}{12}, \forall i = 1, \dots, n$.
- (c) Suponer que se selecciona una muestra de tamaño n mediante muestreo sistemático con probabilidades iguales aplicado sobre la población ordenada de acuerdo a la variable de interés. Obtenga la varianza del estimador de Horvitz-Thompson de la media poblacional en este caso.
- (d) Compare los tres planes de muestreo y justifique cuál resulta más adecuado en la situación planteada.

Observación:

Sea $n \in \mathbb{N}$. Entonces se verifican las siguientes identidades:

- La suma de los primeros n números naturales está dada por:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

- La suma de los cuadrados de los primeros n números naturales está dada por :

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$





Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 2. Un investigador está analizando la duración (en horas) de cierto tipo de bombillas LED. Se asume que la duración de las bombillas sigue una distribución exponencial con parámetro λ . Se toma una muestra aleatoria simple de tamaño $n = 10$, con los siguientes datos (en horas):

$\{88, 102, 96, 91, 110, 85, 99, 107, 95, 100\}$

El fabricante afirma que el tiempo medio de vida es de al menos 100 horas.

- (a) Obtenga en el contexto de este problema, el valor del parámetro λ y qué significado tiene el mismo.
- (b) Formule el contraste de hipótesis para comprobar si los datos contradicen la afirmación del fabricante utilizando un contraste de razón de verosimilitudes.
- (c) Calcule la función de verosimilitud bajo la hipótesis nula y bajo la alternativa (en este caso usa el estimador máximo verosímil)
- (d) Calcule el estadístico de razón de verosimilitudes y determine si se rechaza H_0 al nivel $\alpha = 0,05$.
- (e) Interprete el resultado en el contexto del problema.



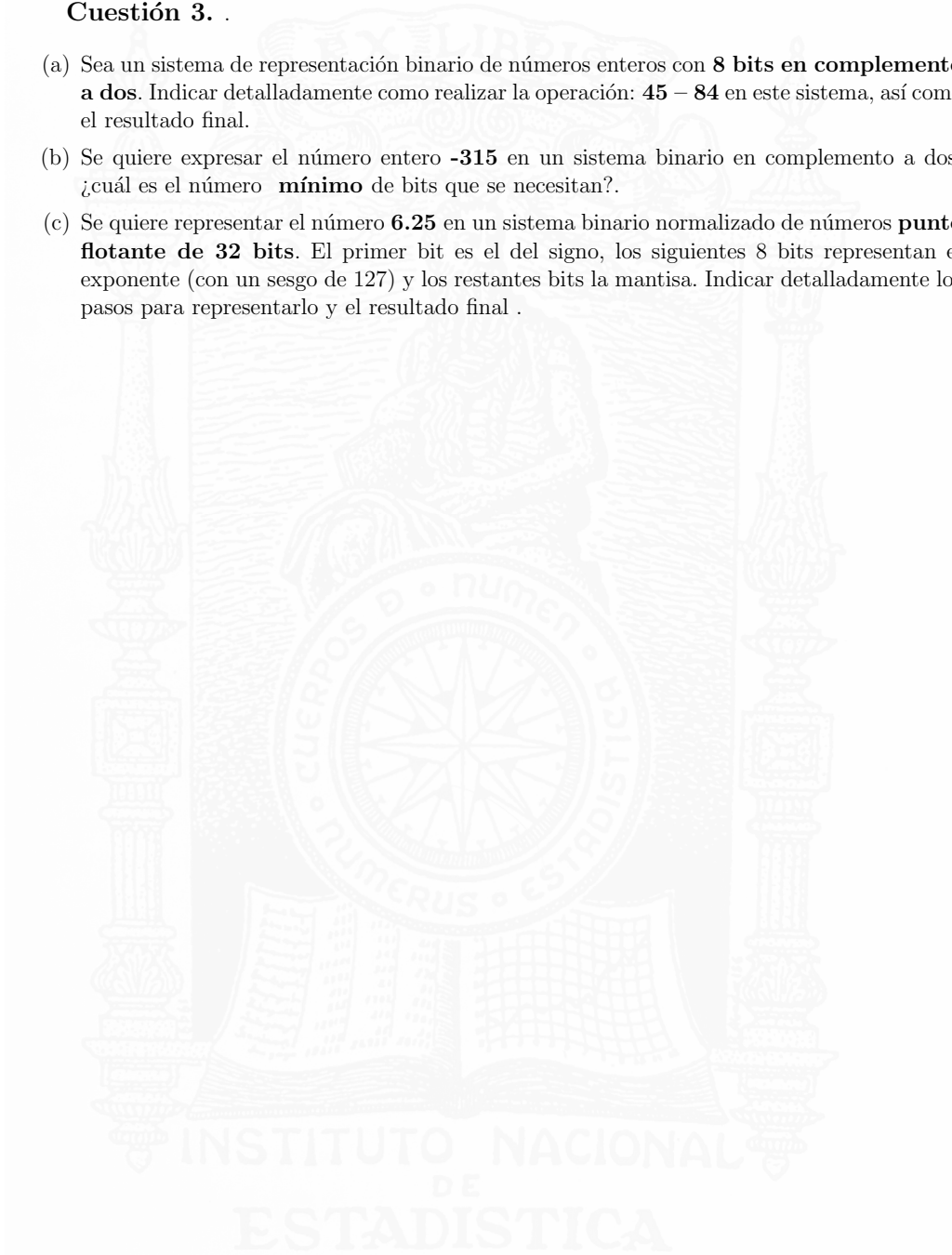
Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 3. .

- (a) Sea un sistema de representación binario de números enteros con **8 bits en complemento a dos**. Indicar detalladamente como realizar la operación: $45 - 84$ en este sistema, así como el resultado final.
- (b) Se quiere expresar el número entero **-315** en un sistema binario en complemento a dos, ¿cuál es el número **mínimo** de bits que se necesitan?.
- (c) Se quiere representar el número **6.25** en un sistema binario normalizado de números **punto flotante de 32 bits**. El primer bit es el del signo, los siguientes 8 bits representan el exponente (con un sesgo de 127) y los restantes bits la mantisa. Indicar detalladamente los pasos para representarlo y el resultado final .





Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 4. Las autoridades municipales desean conocer la utilización de un nuevo tramo del carril bici.

Para ello, se empleó un muestreo por conglomerados sin reemplazamiento y con probabilidades iguales. En concreto, se seleccionaron aleatoriamente 4 semanas del total de 52 que componen el año. Durante cada una de las semanas seleccionadas, se registró diariamente el número de bicicletas que circularon por un punto específico del nuevo tramo del carril bici.

Los datos recogidos se presentan en la siguiente tabla:

Semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	Total Semana
1	20	18	15	17	22	25	30	147
2	12	10	13	11	14	15	17	92
3	14	16	12	13	15	20	21	111
4	15	13	10	12	14	17	18	99

Se pide:

- Una estimación insesgada del total de usuarios anuales del nuevo tramo del carril bici y su error de muestreo relativo o coeficiente de variación. Calcule también un intervalo de confianza del 95 % para dicho total. Considerar $z_{\frac{\alpha}{2}} = 2$.
- Dichas autoridades estarían interesadas en estimar el número medio diario de usuarios del carril bici en un año futuro utilizando el mismo diseño muestral. Con el objetivo de garantizar que la varianza del estimador del número medio de usuarios diarios sea menor que uno, se pide determinar el número mínimo de semanas que deben seleccionarse.



Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 5. Sean los siguientes modelos

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u \quad (\text{Modelo 1})$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + u \quad (\text{Modelo 2})$$

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 \log(x_2) + u \quad (\text{Modelo 3})$$

donde las variables y, x_1, x_2, u son aleatorias tales que los modelos están correctamente definidos.

Se pide:

- (a) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, y obtenga el efecto parcial de x_2 sobre $E(y|x_1, x_2)$ para los modelos 1 y 2.
- (b) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, e interprete el significado de β_1 y β_2 en el modelo 3 (justifique la respuesta).
- (c) Obtenga, para el Modelo 1, $\mathbb{E}(u)$ sabiendo que $\mathbb{E}(u|x_1, x_2) = k$, $k \in \mathbb{R}$
- (d) Considere que $\mathbb{E}(u|x_1, x_2) = 0$, obtenga (si es posible con la información facilitada) y para el Modelo 1 las siguientes covarianzas: $\text{cov}(u, x_1)$, $\text{cov}(u, x_2^2)$.
- (e) Explique el significado y alcance de $\mathbb{E}(u|x_1, x_2) = 0$ en el contexto de la modelización.





Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

Cuestión 6. El ayuntamiento de una ciudad tiene como objetivo estudiar el nivel de actividad de sus comercios locales y para ello ha diseñado una muestra probabilística. Cada comercio seleccionado ha respondido a un cuestionario y finalmente se dispone de los siguientes datos:

Variable	Tipo	Descripción
ID.Comercio	Alfanumérico	Código único de cada comercio. El identificador puede ser cualquier cadena de caracteres que asegure su unicidad, de 4 posiciones
Mes	Entero (1-12)	Mes de referencia de la actividad, donde 1 = Enero, 2 = Febrero, ..., 12 = Diciembre.
Actividad	Binaria (0,1)	Indica si el comercio ha tenido actividad durante el mes (1 = Sí, 0 = No).
Ventas.Tienda (€)	Numérica	Ventas realizadas en el local físico. El valor debe ser 0 o un número positivo.
Ventas.Online (€)	Numérica	Ventas realizadas a través de internet. El valor debe ser 0 o un número positivo.
Ventas.Totales (€)	Numérica	Total de ventas del mes, incluyendo las realizadas en tienda física y online. El valor debe ser 0 o un número positivo.
Gastos (€)	Numérica	Gastos del negocio durante el mes, incluyendo los gastos relacionados con los empleados (salarios, seguros sociales, etc.), alquiler, suministros, entre otros. El valor debe ser 0 o un número positivo.
Empleados	Entero	Número de empleados contratados. El valor debe ser 0 o un número entero positivo.

Se pide:

- (a) Establecer 5 controles de validación (*edits*) para los datos recogidos en la encuesta, cubriendo al menos uno de cada uno de los siguientes tipos: rango, balance y bivalente.

Se debe:

- Escribir la regla de validación como una condición lógica o en pseudocódigo, determinando si es de rango, balance y bivalente.
 - Proporcionar un ejemplo donde se incumpla la regla y explicar brevemente por qué sería un error.
- (b) En los datos recogidos en la encuesta, se ha observado la existencia de valores faltantes en algunas variables de ciertos cuestionarios. Suponiendo que la información recogida corresponde a cinco comercios (registros), se requiere imputar los valores faltantes (NA) utilizando el método indicado y los datos disponibles en la siguiente tabla.



Apéndice B.1. Segundo examen de 2024 del Cuerpo Superior de Estadísticos del Estado.
Promoción interna. Parte común.

Cuerpo Superior de Estadísticos del Estado

Segundo Ejercicio. Convocatoria 2024

ID	Mes	Actividad	Ventas_Tienda	Ventas_Online	Ventas_Totales	Gastos	Empleados
C001	1	1	1200	300	1500	1100	3
C002	1	1	1000	NA	NA	850	2
C003	1	0	0	0	0	0	0
C004	1	1	1200	400	1600	NA	4
C005	1	1	1000	400	1400	1300	5

Métodos de imputación a aplicar:

- **C002 - Ventas_Online:** imputación por la media.
- **C004 - Gastos:** vecino más próximo o cercano.

Indicar para cada caso el valor imputado y los pasos llevados a cabo para obtener dicho valor.

- (c) Suponga que se está valorando la posibilidad de difundir un fichero de microdatos que incluya las variables ventas totales , número de empleados e identificador del comercio para su explotación por parte de usuarios externos. Proponga al menos dos medidas concretas de control del secreto estadístico que podrían aplicarse sobre dichas variables antes de la difusión del fichero.

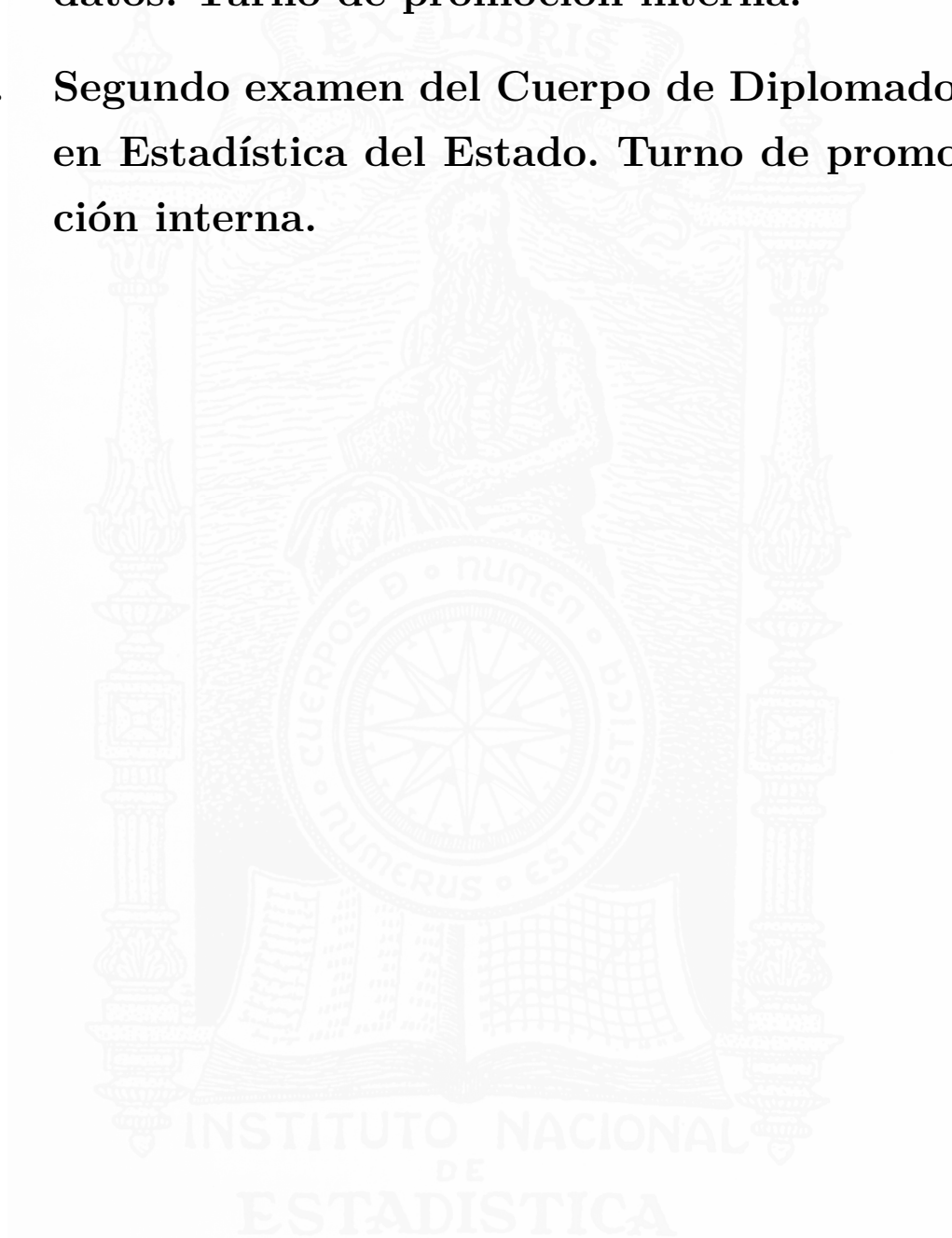




Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Promoción interna.

B.2. Cuarto examen del Cuerpo Superior de Estadísticos del Estado. Especialidad ciencia de datos. Turno de promoción interna.

B.3. Segundo examen del Cuerpo de Diplomados en Estadística del Estado. Turno de promoción interna.





MINISTERIO
DE ECONOMÍA, COMERCIO
Y EMPRESA



**PROCESO SELECTIVO CONVOCADO POR RESOLUCIÓN DE 22 DE
DICIEMBRE DE 2024, DE LA SUBSECRETARÍA DE ECONOMÍA, COMERCIO Y
EMPRESA (BOE 30 DE DICIEMBRE DE 2024)**

**SEGUNDO EJERCICIO DE LA FASE DE OPOSICIÓN DE DIPLOMADOS EN ESTADÍSTICA
DEL ESTADO. PROMOCION INTERNA ESPECIALIDAD ESTADISTICA**

CUESTIONARIO 2025/DEE-PI-ESTADISTICA

AVISO IMPORTANTE PARA LOS OPOSITORES

1. De conformidad con lo establecido en las bases de la convocatoria, deberá elegir y contestar un máximo de **SEIS (6) cuestiones prácticas de las 8 planteadas**. En caso de que exceda el número máximo de cuestiones elegidas, el Tribunal solo corregirá las seis primeras preguntas respondidas, según el orden establecido en el cuestionario del examen, desechando en su caso, el resto de las respuestas. En caso de responder una pregunta que NO quiere que se le corrija, táchela claramente para que el Tribunal lo tenga en cuenta.

2. El tiempo máximo para la realización de este ejercicio es de **130 minutos**.





Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Promoción interna.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado.

CUESTIONES PRACTICAS BASADAS EN EL PROGRAMA DEL ANEXO I

Cuestión práctica 1.

Sean X e Y dos variables aleatorias cuya función conjunta de densidad es la siguiente:

$$f_{XY}(x, y) = \begin{cases} k \cdot x, & 0 < x < y < 2 \\ 0, & \text{resto de casos} \end{cases}$$

Donde k es una constante de valor desconocido. Obtenga razonadamente:

- El valor de k .
- La distribución marginal de Y.
- $P(X < 0,5 / Y = 1)$
- Una aproximación de $P(\sum_{i=1}^{100} W_i < 100)$, siendo $W_1, W_2, \dots, W_{100}$, un total de 100 variables aleatorias independientes con idéntica distribución a X.

Cuestión práctica 2.

En un estudio estadístico se dispone de información de dos variables: X e Y. Se han resumido los datos en la siguiente tabla:

x_i	y_j	n_{ij}
10-30	0-10	3
10-30	10-20	3
30-50	10-20	2
50-110	20-30	2

Además, se han realizado algunos cálculos relativos a los momentos bidimensionales respecto al origen a_{rs} en relación a la citada información (r es el orden para la variable X, y s es el orden para la variable Y):

$$a_{10} = 36 \quad a_{20} = 1.840 \quad a_{01} = 14 \quad a_{02} = 245 \quad a_{11} = 640$$

Se está interesado en realizar predicciones de la variable Y utilizando la información de la variable X. Responda razonadamente a las siguientes preguntas:

- Obtenga la recta de regresión de Y sobre X que minimice la varianza residual.
- Considere las variables Z y W, que se definen como sigue:
 $Z=2X$ $W=Y+3$

Calcule la recta de regresión de W sobre Z que minimice la varianza residual.



Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado.

Cuestión práctica 3.

Un grupo de científicos militares trabajan desarrollando una bomba antes de que lo hagan sus enemigos. Los datos recolectados en 25 pruebas experimentales sugieren que la energía liberada (en kilotones) sigue una distribución normal de media y varianza desconocida.

Las mediciones obtenidas para la media y cuasidesviación típica muestral son las siguientes:

$$\bar{x} = 20,75 \text{ kilotones}$$

$$s_c = 1,25 \text{ kilotones}$$

- Calcule un intervalo de confianza del **95%** para la media de la energía liberada en las explosiones experimentales.
- Se necesita que la energía media liberada supere los **20 kilotones**. Realice un contraste de hipótesis con un nivel de significación del **5%**. ¿Deben los científicos continuar perfeccionando de la bomba, o la energía liberada es suficiente?
- Calcule la región de rechazo para la media, $\bar{x}$, con un nivel de significación de $\alpha = 0,01$. Tenga en cuenta que ahora se sabe que la desviación típica poblacional es $\sigma = 1,5$. La región de rechazo tiene la forma:
$$RC = \{Z > Z_\alpha\}$$
- Si exigimos que el valor superior de la función de potencia sea de **0,0013**, ¿qué media $\bar{x}$ de energía liberada necesitamos alcanzar en **4 pruebas extra** que se vuelven a realizar?

NOTA: Se proporcionan los siguientes valores que pueden ser de utilidad para la resolución de esta cuestión práctica.

Valores tabla t de Student

$$t_{24;0,025} = 2,064; t_{24;0,05} = 1,711; t_{25;0,025} = 2,060; t_{25;0,05} = 1,708$$

Valores tabla Normal

$$Z_{0,05} = 1,64; Z_{0,025} = 1,96; Z_{0,01} = 2,33; Z_{0,0013} = 3$$





Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Promoción interna.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado.

Cuestión práctica 4.

Se dispone de la siguiente información de una economía (en unidades monetarias):

	TOTAL u. m.
Producción	1.700
Consumos intermedios	700
Remuneración de los asalariados	525
Remuneración de los asalariados residentes que trabajan para unidades no residentes	17
Remuneración de los asalariados no residentes que trabajan para unidades residentes	7
Excedente Bruto de Explotación / Renta mixta	425
Impuestos	
- Impuestos sobre los productos	400
- Impuestos sobre la producción y las importaciones	350
Subvenciones	
- Otras subvenciones a los productos	300
- Subvenciones a la producción y las importaciones	200
Rentas netas de la propiedad pagadas al resto del mundo	15
Consumo final efectivo individual	450
Consumo final efectivo colectivo	350

Se pide calcular:

- Valor Añadido Bruto a precios básicos.
- Producto Interior Bruto a precios de mercado de la economía. Debe obtenerse utilizando el Valor Añadido Bruto del apartado anterior.
- Renta Nacional Bruta.
- Gasto en Consumo Final.
- Completar la Cuenta de producción y la Cuenta de explotación de la economía, incluyendo sus saldos.



Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Turno libre.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado.

Cuestión práctica 5.

Considere un monopolista con costes totales $C(q) = 2q + 1$ que se enfrenta a una curva de demanda $P(q) = 10 - q$.

- Halle el equilibrio (precio y cantidad) del monopolio.
- Halle el excedente total del monopolio y la pérdida irrecuperable de eficiencia.
- Halle la función de costes medios e indique los valores de q para que este monopolio sea considerado un monopolio natural.
- La autoridad gubernamental propone una regla de fijación de precios consistente en igualar el precio al coste marginal. Cuantifique cómo varía el excedente total del mercado como consecuencia de dicha política.

Cuestión práctica 6.

En la oficina estadística de un país europeo se dispone de los siguientes datos:

	HOMBRES					MUJERES				
	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años	TOTAL	Menor o igual a 14 años	De 15 a 44 años	De 45 a 64 años	65 y más años
Población media residente en 2024 en miles de personas	9.500	1.900	4.750	1.900	950	10.000	1.800	4.300	2.300	1.600
Defunciones de residentes registradas durante el año 2024	78.755	380	2.375	9.500	66.500	88.370	180	1.290	6.900	80.000

Teniendo en cuenta estos datos, se pide:

- Calcule las tasas específicas de mortalidad por grupos de edad y sexo.
- Usando las tasas del apartado anterior, calcule la tasa bruta de mortalidad por sexo.
- Si tenemos en cuenta las tasas específicas de mortalidad del primer apartado y que la tasa bruta de mortalidad en mujeres es superior a la tasa bruta de mortalidad en hombres, explique cómo se comporta la mortalidad entre ambos sexos. ¿Por qué la mortalidad diferencial por sexo no tiene la misma tendencia en las tasas brutas por sexo que en las tasas específicas por sexo y edad? Explique una mejor forma de hacer la comparativa de las tasas por sexo en este caso.

NOTA: Puede dejar los cálculos de las tasas sin llegar a sus cifras decimales siempre que se indique de forma clara y concisa y cómo ha llegado hasta ellos.





Apéndice B.3. Segundo examen de 2024 del Cuerpo de Diplomados en Estadística del Estado.
Promoción interna.

Ejercicio 2º. Procesos selectivos de ingreso al Cuerpo de Diplomados en Estadística del Estado.

Cuestión práctica 7.

Consideramos los sistemas de representación de números enteros magnitud y signo, complemento a uno y complemento a dos, con 8 bits.

- Explique las diferencias entre estos tres métodos.
- Represente el número entero **-93** con los tres métodos mencionados.
- En un sistema binario (sin bit para signo), ¿A qué valor decimal exacto equivale el número **101010,101**?

Cuestión práctica 8.

El propietario de una finca de naranjos está preocupado por los daños causados por una DANA, y para evaluar dicho desastre decide contratar los servicios de dos estadísticos, A y B, que realizarán un muestreo.

El dueño de la finca les proporciona la siguiente tabla, estratificada según la variedad:

Estratos (variedad)	Número de árboles de la población (N_h)	Proporción estimada de árboles afectados
1	10.000	0,7
2	6.000	0,4
3	4.000	0,15

Si el estadístico A es partidario del muestreo estratificado sin reemplazamiento con afijación proporcional, calcule:

- El tamaño de muestra total que debería utilizar para tener la misma precisión $e = 0,01$ con la que estimaría P en una muestra aleatoria simple sin estratificar.
- El tamaño muestral de cada estrato.

Por otra parte, el estadístico B prefiere el muestreo estratificado con reemplazamiento y afijación óptima. Sin embargo, su material de trabajo lo alquila.

Calcula que, para visitar cada estrato, los costes serán:

$$C_1 = 4 \text{ € por árbol}$$

$$C_2 = 4 \text{ € por árbol}$$

$$C_3 = 9 \text{ € por árbol}$$

- Calcular el tamaño muestral de cada estrato, si el tamaño de muestra total es el mismo que quiere utilizar el estadístico A.



Apéndice C

Galeria de imágenes

C.1. Autoría y referencias bibliográficas de las imágenes portada de cada capítulo

1. Imagen de la portada del capítulo 2 (Año 2024):

Hemos cogido la imagen de la web:

<https://www.google.com/url?sa=i&url=https%3A%2F%2Funiversemagazine.com%2Fen%2F10-years-of-the-historic-jump-from-space-video%2F&psig=AOvVaw2NtAiuJHfPL4KQ4PDEbjW9&ust=1749928990255000&source=images&cd=vfe&opi=89978449&ved=0CBQQjRxqFwoTCNDE9dGP740DFQAAAAAdAAAAABAE> Desconocemos quien es el autor de la fotografía.

2. Imagen de la portada del capítulo 3 (Año 2024):

Hemos cogido la imagen de un artículo de Álex Sala, periodista especializado en Arte e Historia del Arte, en la web:

- https://historia.nationalgeographic.com.es/a/noche-estrellada-maravilloso-paisaje-inventado-por-van-gogh_23406

Desconocemos quien es el autor de la fotografía.





“Numerus regit mundum”
Pitágoras