

Prueba de evaluación continuada 1 - PEC1

Enunciado

- Presentación y objetivos
- Enunciados: descripción teórica de la práctica a realizar
- Materiales
- Criterios de evaluación
- Formato de entrega
- Fecha de entrega

Presentación y objetivos

Objetivos

El objetivo principal de esta **PEC1** es llegar a dominar los conceptos básicos asociados a la estadística descriptiva (medidas de centralización y de dispersión, variables binarias, muestreo,...), así como ser capaz de calcular probabilidades a partir de la distribución normal.

Presentación de la prueba

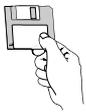
La prueba consta de **tres ejercicios** más **dos prácticas** de Minitab.

En cada **ejercicio**, el alumno tendrá que resolver y responder razonadamente las preguntas que se formulan sin recurrir al Minitab, sólo trabajando con tablas y calculadora.

Los **ejercicios prácticos**, se resolverán con la ayuda del programa Minitab. Por ello no solo tenéis que enviar el fichero Word con los ejercicios desarrollados, sino también el ficheros de Minitab con extensión ***.MPJ**.

Enunciado de la PEC1

Parte 1 de la Práctica: Preguntas



EJERCICIO 1

Se supone que el 2.50% de los centros de investigación de Cataluña tienen más de 80 trabajadores y que un 28.10% de los centros de investigación tienen menos de 40 trabajadores. Si suponemos que el número de trabajadores de los centros de investigación catalanas sigue una distribución normal,

Se pide:

 **E1.1.** La media (μ) y la desviación típica (σ) de esta distribución.

Definimos $X = \text{"Número de trabajadores"}$.

$$P(X > 80) = 0.025 \Rightarrow P(X \leq 80) = 0.975 \Rightarrow P\left(Z \leq \frac{80 - \mu}{\sigma}\right) = 0.975 \Rightarrow \frac{80 - \mu}{\sigma} = 1.96$$

$$P(X \leq 40) = 0.281 \Rightarrow P\left(Z \leq \frac{40 - \mu}{\sigma}\right) = 0.281 \Rightarrow \frac{40 - \mu}{\sigma} = -0.58$$

Tenemos dos ecuaciones y dos incógnitas a partir de las cuales encontramos:

$$\sigma = 15.75 \quad \text{y} \quad \mu = 49.13$$

 **E1.2.** La probabilidad de que un centro de investigación escogido al azar tenga más de 20 trabajadores.

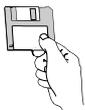
$$P(X > 20) = 1 - P(X \leq 20) = 1 - P\left(Z \leq \frac{20 - 49.13}{15.75}\right) = 1 - P(Z \leq -1.84) =$$

$$= 1 - 0.03288 = 0.967$$

 **E1.3.** La probabilidad de que la media de trabajadores de 10 centros sea inferior a 40.

Definimos $Y = \text{"Media de trabajadores"}$.

$$P(Y \leq 40) = P\left(Z \leq \frac{40 - 49.13}{15.75/\sqrt{10}}\right) = P\left(Z \leq \frac{-9.13}{4.98}\right) = P(Z \leq -1.83) = 0.03362$$



EJERCICIO 2

Se ha hecho una evaluación de las bibliotecas de una ciudad analizando, por una parte los fondos de que disponían y, por otra, la accesibilidad. Los resultados a los que se llegó son los siguientes.

$$\bar{X}_F = 15'5 \qquad \bar{X}_A = 75$$

$$\sigma_F = 2'5 \qquad \sigma_A = 30'6$$

Si dos bibliotecas (nombradas B_A y B_B) han obtenido las siguientes puntuaciones:

$$B_A \text{ fondos} = 16'7 \quad B_A \text{ accesib} = 77'5$$

$$B_B \text{ fondos} = 14 \quad B_B \text{ accesib} = 82'4$$

Se pide:

 **E2.1.** Convertir las puntuaciones de los dos apartados (fondos y accesibilidad) de las dos bibliotecas y decir cuál de las dos ha obtenido una valoración global más favorable.

Lo que se tiene que hacer es encontrar las medias de los dos apartados para cada biblioteca, sumarlas y compararlas. Pero para esto, para que se puedan sumar y comparar, primero las tenemos que estandarizar.

$$B_A \text{ fondos} = (16'7 - 15'5) / 2'5 = 0'48 \qquad \text{i} \qquad B_A \text{ accesib} = (77'5 - 75) / 30'6 = 0'08.$$

$$B_B \text{ fondos} = (14 - 15'5) / 2'5 = -0'60 \qquad \text{i} \qquad B_B \text{ accesib} = (82'4 - 75) / 30'6 = 0'24$$

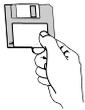
0'24.

Si ahora hacemos la media, tenemos:

$$BA = (0'48 + 0'08)/2 = 0'28$$

$$BB = (-0'60 + 0'24)/2 = 0'18$$

Por lo tanto, $BA > BB$, la biblioteca A tiene una mejor valoración global.



EJERCICIO 3

En una compañía de seguros el departamento de inspección de siniestros ha hecho una valoración de los daños ocasionados por cierto tipo de accidentes y una estimación de las causas que los produjeron, y los resultados se muestran en la tabla siguiente:

| Valoración de los daños (en miles de euros) | Número de accidentes provocados | Número de accidentes no provocados |
|---|---------------------------------|------------------------------------|
| 20-40 | 2 | 68 |
| 40-80 | 12 | 188 |
| 80-120 | 18 | 222 |
| 120-160 | 10 | 210 |
| 160-200 | 2 | 98 |
| 200-240 | 1 | 69 |
| 240-280 | 1 | 59 |
| 280-320 | 1 | 29 |
| 320-400 | 0 | 10 |

Se pide:



E3.1.

Encontrar la media de los daños valorados por accidente.

Se trataría simplemente de buscar la **media** aritmética de la primera

columna: $\bar{x} = \sum_{i=1}^{i=k} x_i f_i = 130.500/1000 = 130'5$.



E3.2.

Encontrar el valor del daño provocado más frecuente.

Aquí se trataría de buscar la **moda** en la segunda columna:

$$Mo = L_{i-1} + \frac{d_1}{d_1 + d_2} \cdot a_i = 80 + [6/(8+6)] \cdot 40 = 97'142857, \text{ pero lo de multiplicar}$$

por 1000 y tendremos: 97.142'86€.



E3.3.

Encontrar el valor de la mediana de los daños no provocados.

Aquí se trataría de encontrar la **mediana** de la tercera columna, como $N/2 = 476'5$, la mediana estará en el intervalo (80-120):

$$Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} \cdot c_i = 80 + \frac{[476'5 - 256]}{222} \cdot 40 = 119'72972, \text{ que al multiplicar por } 1000 \text{ tenemos: } 119.729'72\text{€.}$$

Se pide:

Parte 2 de la Práctica: Prácticas Minitab



PRÁCTICA 1

La biblioteca de una prestigiosa universidad española, quiere realizar un estudio sobre los préstamos realizados por cada facultad al cabo del año. Para eso, lo que hacemos es extraer una muestra de 150 observaciones para cada uno de las facultades. En principio, haremos la prueba para los estudios de 'Empresariales', 'Biblioteconomía', y 'Sociología'. Los datos obtenidos los encontraréis en el documento **prest.mtw**.

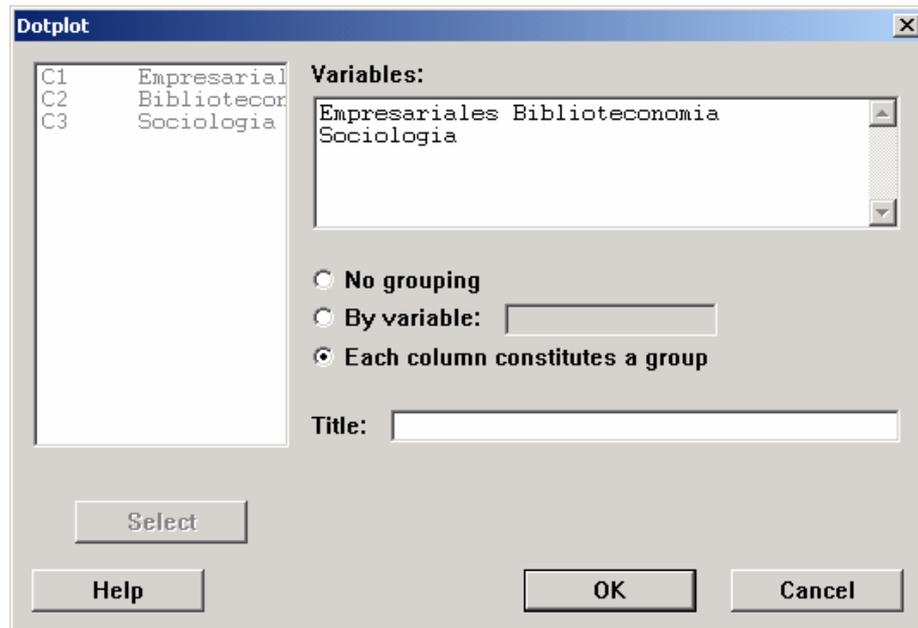
Se pide:



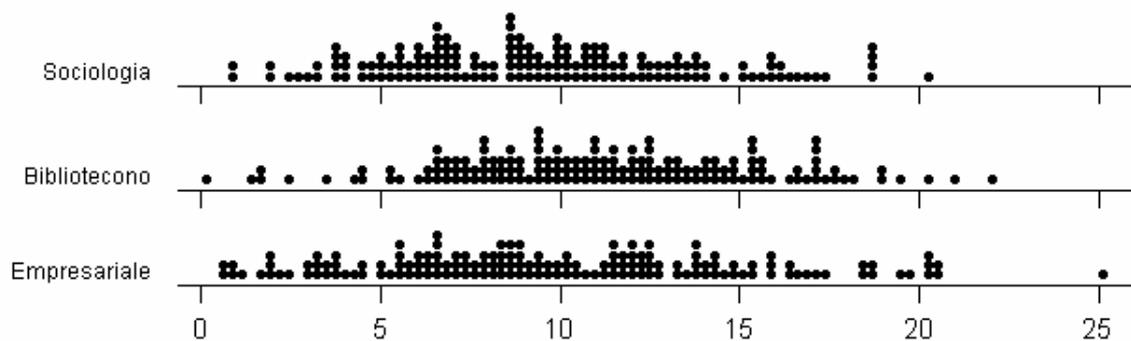
P1.1.

Realizar un *dotplot* sobre cada uno de los estudios y comentar los resultados.

Para dibujar el dotplot correspondiente a cada grupo, seleccionamos *Graph > Dotplot:*



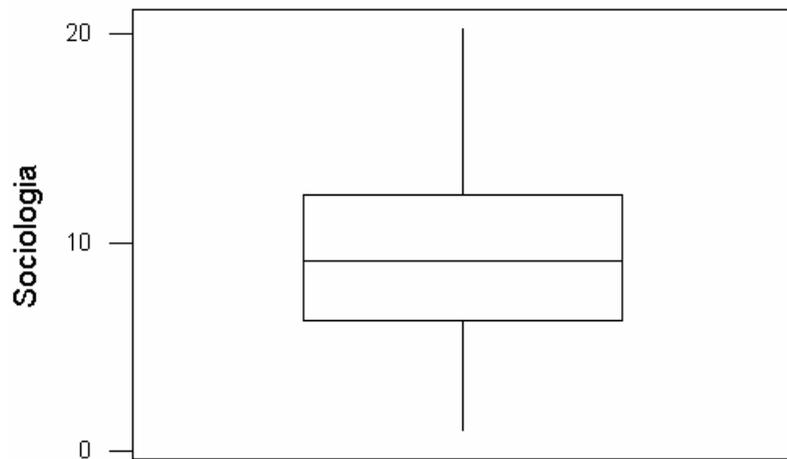
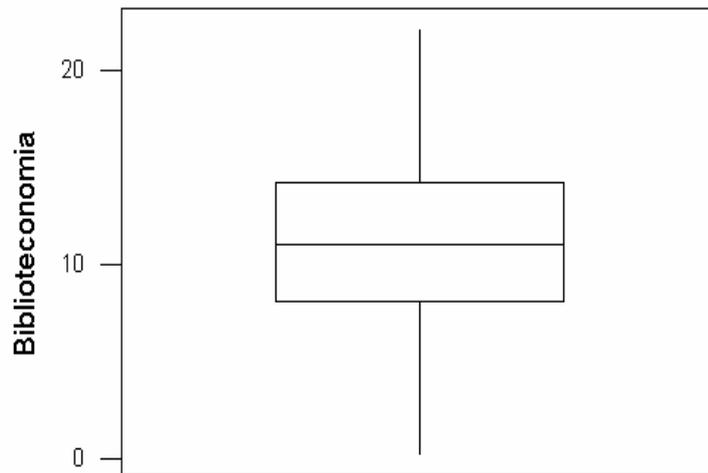
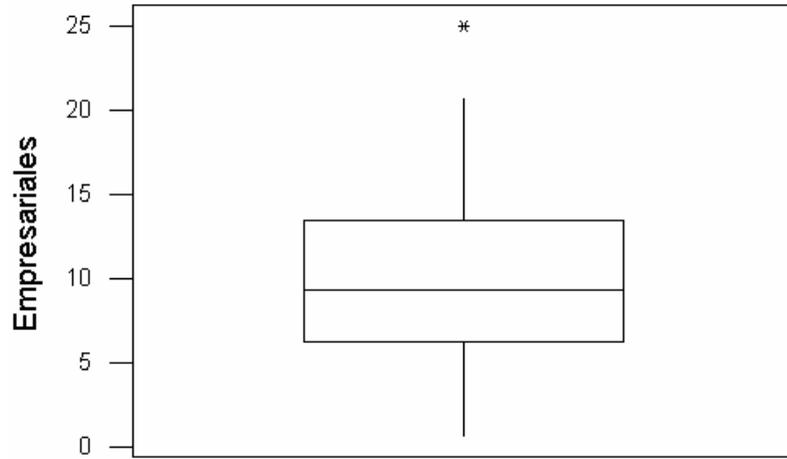
Dotplot for Empresariales-Sociologia



De los gráficos anteriores podríamos razonar que los tres siguen una distribución aprox. Normal. Parece que sean los estudiantes de Sociología y los de Biblioteconomía los que tienden a realizar un mayor número de préstamos anuales.

P1.2. Dibujar un diagrama de cajas (*boxplot*) asociado a cada estudio, así como también los estadísticos descriptivos correspondientes.

Para realizar el diagrama de cajas, seleccionamos *Graph > Boxplot*, y en el eje de las Y, insertamos cada una de las columnas:



Para calcular los estadísticos descriptivos asociados a cada estudio, seleccionamos *Stat > Basic Statistics > Display Basic Statistics*:

Descriptive Statistics: Empresariales; Biblioteconomía; Sociología

| Variable | N | Mean | Median | TrMean | StDev | SE Mean |
|----------|-----|--------|--------|--------|-------|---------|
| Empresar | 150 | 9,835 | 9,265 | 9,701 | 5,112 | 0,417 |
| Bibliote | 150 | 11,204 | 11,042 | 11,224 | 4,232 | 0,346 |
| Sociolog | 150 | 9,448 | 9,158 | 9,354 | 4,189 | 0,342 |

| Variable | Minimum | Maximum | Q1 | Q3 |
|----------|---------|---------|-------|--------|
| Empresar | 0,558 | 25,023 | 6,159 | 13,422 |
| Bibliote | 0,223 | 22,087 | 8,101 | 14,232 |
| Sociolog | 1,000 | 20,215 | 6,279 | 12,263 |

A partir de estos datos comprobamos que, efectivamente, los estudios de Biblioteconomía son los que tienen la media de préstamos más alta respecto al resto de los estudios. Mientras que los estudios de Empresariales y Sociología tienen una media parecida.

 **P1.3.** El sesgo de una distribución depende de la simetría de ésta. A partir de los datos del anterior apartado, explica qué tipo de asimetría poseen las distribuciones correspondientes a 'Empresariales', 'Biblioteconomía', y 'Sociología'.

-  **Empresariales:** Distribución sesgada hacia la derecha porque la media es superior a la mediana
-  **Biblioteconomía:** Distribución ligeramente sesgada hacia la derecha porque la media es ligeramente superior a la mediana. En este caso, casi tenemos una distribución simétrica.
-  **Sociología:** Análogamente casos anteriores. Tenemos una distribución ligeramente sesgada a la derecha.



PRÁCTICA 2

Uno de los aspectos que se pretende mejorar en la biblioteca del apartado anterior es la reducción de tiempo de espera de cada estudiante a la hora de realizar un préstamo. Evidentemente, el tiempo de espera depende de muchos factores como la hora del día, época de exámenes, etc. En estos momentos, se sabe que el tiempo medio de espera es de 6 minutos con una desviación estándar de 4.7 minutos.

Si tomamos al azar un grupo de 70 estudiantes:

Se pide:



¿Cuál es la probabilidad de que el tiempo medio de espera sea

P2.1. menor de 5 minutos?

Como $n=35 > 30$, podremos utilizar el Teorema Central del Límite, para afirmar que la distribución de las medias muestrales \bar{X} se podrá aproximar por una normal con media 6 y desviación estándar 0.56, es decir,

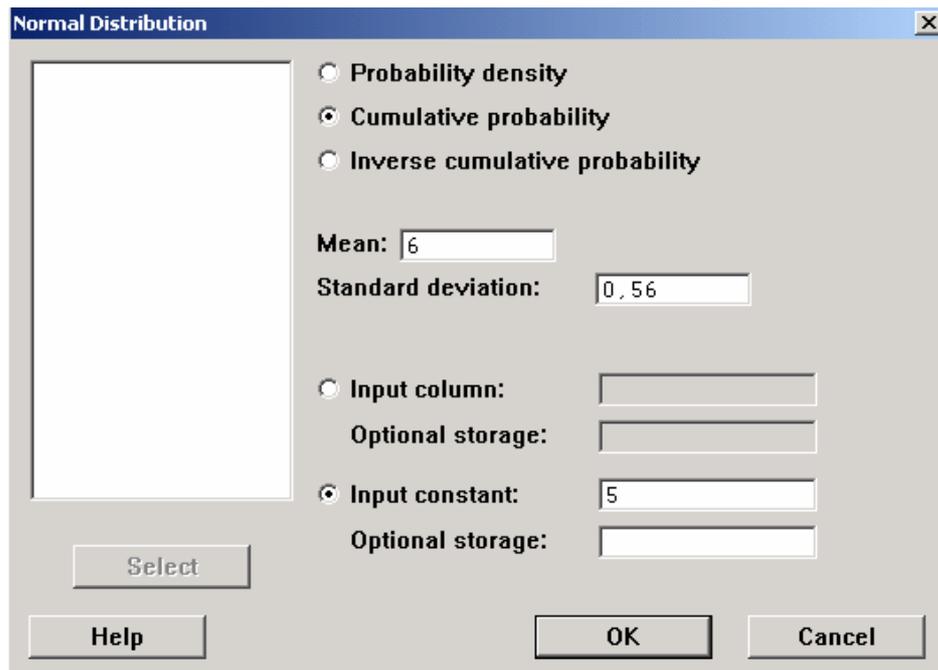
$X =$ "Tiempo de espera de un estudiante en la cola de una biblioteca"

$\bar{X} =$ "Tiempo medio de espera de un grupo de 70 estudiantes"

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \sim N\left(6, \frac{4.7}{\sqrt{70}}\right)$$

Por tanto, lo que queremos calcular es $P(\bar{X} < 5)$

Seleccionamos *Calc > Probability Distributions > Normal*, y activamos la opción *Cumulative Probability*:



Cumulative Distribution Function

Normal with mean = 6,00000 and standard deviation = 0,560000

| x | P(X <= x) |
|--------|-------------|
| 5,0000 | 0,0371 |

Por tanto, la probabilidad de que el tiempo medio de espera sea inferior a 5 minutos es de un 3,71 %

 **P2.2.** ¿Cuál es la probabilidad que el tiempo medio esté entre 4 y 8 minutos?

Queremos calcular $P(4 < \bar{X} < 8) = P(\bar{X} < 8) - P(\bar{X} < 4)$

De forma análoga al apartado anterior calculamos las dos probabilidades:

Cumulative Distribution Function

Normal with mean = 6,00000 and standard deviation = 0,560000

| x | P(X <= x) |
|--------|-------------|
| 4,0000 | 0,0002 |

Cumulative Distribution Function

Normal with mean = 6,00000 and standard deviation = 0,560000

| x | P(X <= x) |
|--------|-------------|
| 8,0000 | 0,9998 |

Por tanto, $P(4 < \bar{X} < 8) = 0.9998 - 0.0002 = 0.9996$.

Es decir, la probabilidad de que el tiempo de espera esté entre 4 y 8 minutos es prácticamente del 100%.

 **P2.3.** Si queremos conseguir una probabilidad exacta del 0.90, ¿qué tiempo medio de espera obtendríamos?

Queremos conocer t, tal que $P(\bar{X} < t) = 0.90$

Para ello, seleccionamos *Calc > Probability Distributions > Normal*, activando *Inverse Cumulative Probability*:

Inverse Cumulative Distribution Function

Normal with mean = 6,00000 and standard deviation = 0,560000

| P(X <= x) | x |
|-------------|--------|
| 0,9000 | 6,7177 |

Por tanto, para conseguir una probabilidad del 0.90 el tiempo medio estimado sería aproximadamente de 6,7 minutos.



Materiales

Para trabajar esta **PEC1**, se necesita haber trabajado:

- Los módulos 1 al 11
- Las **GES1**, **GES2** y **GES3**
- Los vídeos correspondientes
- Los materiales complementarios enviados al [Tablón](#)
- El software Minitab: especialmente los [documentos de actividades](#) resueltas



Criterios de evaluación

La práctica se evaluará atendiendo a los siguientes criterios:

- Los **ejercicios** se puntuarán con un máximo de 2 puntos.
- Cada **práctica** de Minitab se puntuará con un máximo de 2 puntos.
- La puntuación máxima se consigue cuando la respuesta y su justificación son correctas.



Formado de entrega

La **PEC1** se entregará en un documento **Word** para Windows en el buzón de: "Ev.Continua" que tenéis en el aula junto con el [Tablón](#) y [Foro](#).

El nombre del fichero/documento tendrá el siguiente formado: "ape1_ape2_PEC1.doc". Los apellidos se escribirán sin acentos. Por ejemplo, el archivo de un estudiante nombrado Alfredo García Melgar llevaría la siguiente denominación: garcia_melgar_PEC1.doc".

Cada alumno tendrá que enviar dos archivos:

- **Garcia_melgar_PEC1.doc**
- **Garcia-melgar_PEC1.MPJ**

*** * *** MUY IMPORTANTE: Para poder entender las fórmulas, nos tendríamos que acostumbrar a hacerlas con el Editor de ecuaciones del programa Word.



Fecha de entrega

La **fecha límite de entrega** son las **24 horas** del día **3 de abril del 2005**. Recordad que las fechas límite tienen que respetarse estrictamente.

Asimismo, el día 4/03/2005, a primera hora se publicará la resolución de la **PEC1**.

