ESCUELA SUPERIOR DE INGENIERÍA
GRADO INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES — ESTADÍSTICA (3-SEPTIEMBRE-2012)

Nombre y Apellidos:

EJERCICIOS

1. **Un operario de una fábrica de neumáticos inicia la comprobación de rutina del sensor de verificación de la soldadura de la napa carcasa en un puesto de montaje. Los resultados obtenidos sobre un total de 35 muestras son los siguientes:**

Soldadura	**Grosor (mm)**	f_i
Déficit	**[0 - 3.5)**	**0.11**
Normal	**[3.5 - 6.5)**	**0.65**
Exceso	**[6.5 - 8.5)**	
Inadmisible	**[8.5 - 10)**	**0.08**

a) **(0.25 ptos.) ¿Qué porcentaje de los datos tuvo una soldadura en exceso o inadmisible?**

b) **(0.50 ptos.) La normativa de neumáticos especifica que si el 30 % de los datos supera los 5 mm de grosor de soldadura, el sensor debe ser llevado a revisión para un recalibrado ¿Es necesario efectuar dicha revisión? ¿Por qué?**

c) **(0.25 ptos.) Obtener el coeficiente de variación de los datos.**

SOLUCIÓN: Completamos la tabla de frecuencias sabiendo que la suma de frecuencias relativas debe ser igual a 1. También incluimos la marca de clase de cada intervalo, el resto de frecuencias y los cálculos necesarios para el calculo de la media y la desviación típica de los datos.

Soldadura	**Grosor (mm)**	x_i	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
Déficit	$[0-3{,}5)$	1,75	0,11	0,11	0,1925	0,336875
Norma	$[3{,}5-6{,}5)$	5,00	0,65	0,76	3,25	16,25
Exceso	$[6{,}5-8{,}5)$	7,50	0,16	0,92	1,2	9
Inadmisible	$[8{,}5-10)$	9,25	0,08	1	0,74	6,845

a) Sumando las frecuencias relativas que corresponden a las categorías “en exceso” e “inadmisible” obtenemos $0{,}16 + 0{,}08 = 0{,}24$ que representa al $\boxed{24\,\%}$ de los datos.

b) Para determinar el grosor que superan el 30 % de los datos es necesario calcular el percentil 70 de esta distribución. Para ello, observamos en la tabla el primer intervalo con frecuencia relativa acumulada que supera el valor 0,7, en este caso el intervalo $[3{,}5-6{,}5)$, e interpolamos un dato dentro de este intervalo:

$$\overbrace{3{,}5\ \overbrace{\qquad\qquad}^{x}\ P_{70}\qquad 6{,}5}^{3}$$
$$\underbrace{0{,}11\ \underbrace{\qquad\qquad}_{0{,}59}\ 0{,}7\qquad 0{,}76}_{0{,}65}$$

$$\begin{array}{lcl} 0{,}65 & \longrightarrow & 3 \\ 0{,}59 & \longrightarrow & x \end{array}$$

luego

$$x = \frac{0{,}59 \cdot 3}{0{,}65} = 2{,}723 \quad \Longrightarrow \quad P_{70} = 3{,}5 + x = \boxed{6{,}223\ mm}$$

De este modo hemos determinado que el 30 % de los datos superan los 6,223 mm y por tanto se deduce que el porcentaje de datos que superan los 5 mm es mayor al 30 %. Debemos por tanto realizar una revisión de la maquinaria.

c) Los cálculos necesarios para determinar el coeficiente de variación son:

$$\overline{x} = \sum x_i \cdot f_i = 5{,}3825 \qquad S^2 = \sum x_i^2 \cdot f_i - \overline{x}^2 = 32{,}231875 - 5{,}3825^2 = 3{,}2606$$

El coeficiente de variación se calcula como:

$$CV = \frac{S}{|\overline{x}|} = \frac{1{,}8057}{5{,}3825} = \boxed{0{,}3355}$$

NOTA: El ejercicio 1 también puede resolverse utilizando frecuencias absolutas en lugar de frecuencias relativas. Para ello en la tabla inicial de cálculos debemos multiplicar las frecuencias relativas por el número de datos (en este caso 35) y obtener las correspondientes frecuencias absolutas (n_i) de cada intervalo. El razonamiento con frecuencias absolutas es el mismo.

2. **Un ex alumno de GITI de la ESI de Cádiz está trabajando en la empresa Krauss-Maffei-Wegmann en Munich fabricando motores para trenes TALGO de última generación. En su control de calidad descubre que un error en una máquina está provocando que el 40 % de los motores produzcan unos valores altos de dióxido de carbono (CO_2), el 35 % unos niveles altos de dióxido de azufre (SO_2) y el 5 % unos valores altos de ambos subproductos de combustión (SO_2 y CO_2). Para poder aislar el problema, su jefe de sección, el señor Zimmerman, le solicita un informe que incluya la siguiente información:**

   ***a*) (0.25 ptos.) La probabilidad de que un motor que muestre niveles altos de dióxido de carbono CO_2, tenga también niveles altos de dióxido de azufre SO_2.**

   ***b*) (0.25 ptos.) La probabilidad de que un motor que muestre niveles altos de dióxido de azufre SO_2, no tenga niveles altos de monóxido de carbono CO_2.**

   ***c*) (0.25 ptos.) La probabilidad de que un motor tenga valores normales de ambos gases.**

   ***d*) (0.25 ptos.) La probabilidad de que un motor tenga emisión elevada de gases.**

SOLUCIÓN: Denominamos a los sucesos del siguiente modo:

$$CO_2 = \{\text{El motor produce valores altos de dióxido de carbono}\}$$
$$SO_2 = \{\text{El motor produce valores altos de dióxido de azufre}\}$$

Seguidamente asignamos a los diferentes sucesos que aparecen en el enunciado del ejercicio sus correspondientes probabilidades:

$$P(CO_2) = 0{,}40 \qquad P(SO_2) = 0{,}35 \qquad P(CO_2 \cap SO_2) = 0{,}05$$

Utilizando fórmulas elementales de probabilidad damos respuesta a los apartados, teniendo en cuenta que para responder al apartado(c) hemos resuelto antes el apartado (d):

a) $P(SO_2/CO_2) = \frac{P(SO_2 \cap CO_2)}{P(CO_2)} = \frac{0{,}05}{0{,}40} = \boxed{0{,}125}$

b) $P\left(\overline{CO_2}/SO_2\right) = 1 - P(CO_2/SO_2) = 1 - \frac{P(CO_2 \cap SO_2)}{P(SO_2)} 1 - \frac{0{,}05}{0{,}35} = \boxed{0{,}857}$

c) $P\left(\overline{CO_2} \cap \overline{SO_2}\right) = P\left(\overline{CO_2 \cup SO_2}\right) = 1 - P(CO_2 \cup SO_2) = 1 - 0{,}7 = \boxed{0{,}3}$

d) $P(CO_2 \cup SO_2) = P(CO_2) + P(SO_2) - P(CO_2 \cap SO_2) = 0{,}40 + 0{,}35 - 0 - 05 = \boxed{0{,}7}$

3. El número medio de clientes que llegan a un supermercado es de 0.1 clientes por minuto.

   ***a*) (0.5 ptos.) ¿Cuál es la probabilidad de que lleguen al menos 5 clientes en una hora?**

   ***b*) (0.5 ptos.) Se observa que en este mismo supermercado, en la jornada del sábado por la tarde (de 17:00h a 22:00h) llegan 400 clientes en término medio. Calcula la probabilidad de que en dicha jornada haya más de 450 clientes.**

SOLUCIÓN:

a) La variable aleatoria X =*"Nº de usuarios que llegan al supermercado en 1 hora"* sigue una distribución de Poisson de parámetro $\lambda = 6$

$$X \sim Po(6) \quad \Rightarrow \quad P[X \geq 5] = 1 - P[X < 5] = 1 - (p_0 + p_1 + p_2 + p_3 + p_4) =$$

$$= 1 - e^{-6}\left(\frac{6^0}{0!} + \frac{6^1}{1!} + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!}\right) = \boxed{0.7149}$$

b) La variable aleatoria Y =*"Nº de usuarios que llegan al supermercado un sábado cualquiera por la tarde"* sigue una distribución de Poisson de parámetro $\lambda = 400$. Esta distribución puede aproximarse por la Normal (se verifica que $np, nq > 5$ y $p, q > 0{,}05$):

$$X \sim Po(400) \simeq N(\mu;\sigma) \quad \begin{cases} \mu = \lambda = 400 \\ \sigma = \sqrt{\lambda} = \sqrt{400} = 20 \end{cases}$$

Realizando la corrección por continuidad calculamos:

$$P[X > 450] \overset{(c.c.)}{=} P[X > 450{,}5] = P[Z > 2{,}525] = 1 - F_Z(2{,}525) = 1 - 0{,}9943 = \boxed{0{,}0057}$$

4. Los siguientes datos representan el tiempo medio (en minutos) de duración de las películas que producen dos compañías cinematográficas:

Compañía 1	**103**	**94**	**110**	**87**	**98**
Compañía 2	**97**	**82**	**123**	**92**	**175**

***a*) (0.5 ptos.) Construye un intervalo de confianza al 95 % para el cociente de varianzas.**

***b*) (0.5 ptos.) Constrasta con un nivel de significación del 5 % si podemos considerar que el tiempo medio de duración de las películas de ambas compañías es el mismo.**

SOLUCIÓN: A partir de los datos del enunciado determinamos los principales parámetros muestrales de ambas muestras (utilizamos el modo estadístico de nuestra calculadora):

Compañía 1	Compañía 2
$\overline{x}_1 = 98{,}4$	$\overline{x}_2 = 113{,}8$
$S^2_{c_1} = 76{,}3$	$S^2_{c_2} = 1399,7$
$n_1 = 5$	$n_2 = 5$

a) Estimamos el cociente entre las varianzas poblaciones con un intervalo de confianza al 95 %:

$$Idc_{1-\alpha}\left[\frac{\sigma_1^2}{\sigma_2^2}\right] = \left(\frac{S^2_{c_1}/S^2_{c_2}}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)} \,;\, \frac{S^2_{c_1}/S^2_{c_2}}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)}\right)$$

$$= \left(\frac{76{,}3/1399{,}7}{F_{0{,}975}(4,4)} \,;\, \frac{76{,}3/1399{,}7}{F_{0{,}025}(4,4)}\right) = \left(\frac{0{,}0545}{9{,}605} \,;\, \frac{0{,}0545}{0{,}104}\right) = \boxed{(\ 0.006\ ;\ 0.524\ )}$$

siendo el percentil $F_{0{,}975}(4,4) = 9{,}605$, según la tabla, y utilizando que

$$F_{\frac{\alpha}{2}}(n_1-1,n_2-1) = \frac{1}{F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)} \quad \Rightarrow \quad F_{0{,}025}(4,4) = \frac{1}{F_{0{,}975}(4,4)} = \frac{1}{9{,}605} = 0{,}104$$

b) Para analizar si existen diferencias significativas en la duración media de las películas de las 2 compañías, suponiendo Normalidad en las poblaciones, deberíamos contrastar primero la igualdad de varianzas. Sin embargo en este caso no es necesario, porque el intervalo de confianza que hemos calculado en el apartado (a) no contiene el valor 1, por lo que se deduce que las varianzas poblaciones no pueden ser iguales. Por tanto utilizaremos el contraste de comparación de medias suponiendo varianzas distintas:

$$\left.\begin{array}{l} H_0 \equiv \mu_1 = \mu_2 \\ H_1 \equiv \mu_1 \neq \mu_2 \end{array}\right\} \qquad t_{exp} = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{S^2_{c_1}}{n_1} + \frac{S^2_{c_2}}{n_2}}} = \frac{98{,}4 - 113{,}8}{\sqrt{\frac{76{,}3}{5} + \frac{1399{,}7}{5}}} = -0{,}8963$$

En este caso el estadístico t_{exp} se ajusta a una distribución $t(g)$ donde

$$g = \frac{\left(\frac{S_{c_1}^2}{n_1} + \frac{S_{c_2}^2}{n_2}\right)^2}{\frac{\left(\frac{S_{c_1}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_{c_2}^2}{n_2}\right)^2}{n_2 - 1}} = 4{,}4348 \simeq 4$$

La región crítica del contraste al 5 % de significación es

$$\text{R.C.} = \{|t_{exp}| > t_{0{,}975}(4)\} = \{|t_{exp}| > 2{,}776\} \quad \Rightarrow \quad H_0,$$

$\boxed{\text{Por tanto no podemos rechazar } H_0 \text{ para } \alpha = 0{,}05\text{, es decir, no podemos afirmar que existen diferencias significativas entre la duración de las películas de ambas compañías.}}$