

1. (1 punto) Sean A y B dos sucesos asociados a un mismo experimento aleatorio con $P(A) = 0'7$, $P(B) = 0'6$ y $P(\bar{A} \cup \bar{B}) = 0'58$.
 - a. ¿Son A y B independientes?
 - b. Si $E \subset A$, calcula $P(\bar{E}/\bar{A})$.


2. (1 punto) Sean X e Y dos variables aleatoria, independientes y consideremos una nueva variable $Z = 2X - 3Y + 6$. Calcula el valor esperado y la varianza de Z suponiendo que X sigue una distribución de Poisson con frecuencia $\lambda = 0'3$ e Y es una variable normal de media 3 y varianza 4.

3. (1'5 puntos) Un sistema de comunicación binario transmite solo 0 y 1. Debido al ruido del sistema, a veces un 0 transmitido se recibe como un 1, y un 1 transmitido se recibe como un 0. Suponiendo que la probabilidad de que un 0 se transmite correctamente es 0'94, que la probabilidad de que un 1 se transmita correctamente es 0'91 y que la proporción de ceros transmitidos es 0'45, calcular:
 - a. La probabilidad de que habiendo recibido 1, haya sido transmitido un 1.
 - b. La probabilidad de que habiendo recibido un 0, haya sido transmitido un 1.

4. (1'75 puntos) A la Biblioteca Universitaria Concepción Arenal en el Campus Vida llegan, en promedio, 9 estudiantes cada 15 minutos. Calcula:
 - a. La probabilidad de que en una hora lleguen más de 30 estudiantes.
 - b. La probabilidad de que el tiempo entre la llegada de 2 estudiantes consecutivos sea menor de 2 minutos.
 - c. El tiempo medio que hay que aguardar hasta que lleguen 60 estudiantes.

Nota: En este ejercicio puedes utilizar comandos de R para responder, pero solo aquellos relacionados con la normal estándar (media 0 y varianza 1).


5. (1'75 puntos) En el incendio del cementerio de neumáticos de Seseña (Toledo) se han registrado las mediciones de emisiones contaminantes en 45 zonas, calificando contaminantes 12 zonas.
 - a. Obtén un intervalo de confianza para la proporción de zonas contaminadas, con un nivel del 99%.
 - b. Cuando el porcentaje de zonas contaminadas supera el 30%, se produce una actuación de los servicios medioambientales. Con un nivel de significación del 5%, ¿dirías que es necesaria una intervención)?

 Escola Técnica Superior de Enxeñaría	Grado en Ingeniería Informática - Universidade de Santiago de Compostela	
	Asignatura: Estadística	Curso: 2021-2022
	Nombre:	Prueba Ev. Cont. 1
	Apellidos:	Nota:

- Del siguiente conjunto de datos de la variable X , $\{x_i, i = 1, \dots, 6\} = \{2, 3, 4, 4, 5, 6\}$.
 - ¿Podría ser la media menor que 2?
 - Calcula la media aritmética, la moda y la mediana.
 - Calcula la varianza, la desviación típica y el rango muestrales.
 - Calcula la media de los valores transformados $y_i = 3x_i - 2$.
- Unos conocidos productores de café de Colombia utilizan compañías aéreas locales para enviar el café producido desde las montañas al aeropuerto internacional más cercano. Por razones de coste, el 65 % de las veces contratan a la compañía AirWings, mientras que los viajes restantes los realizan con LifeFlight. Ambas compañías poseen aviones Tupolev (la mitad de las aeronaves de AirWings y el 75 % de las de LifeFlight son de este fabricante). Calcula:
 - La probabilidad de que uno de los envíos no se realice en Tupolev.
 - Si el envío desde las montañas ha sido realizado en un Tupolev, calcula la probabilidad de que la compañía que lo ha transportado sea LifeFlight.
 - La probabilidad de que el envío sea con AirWings o en Tupolev.
- En la primera parte del examen de Estadística del curso 2019-2020 se hizo una prueba tipo test de 20 preguntas. Un alumno que ha preparado la materia concienzudamente, tiene un 80 % de posibilidades de responder bien a cada pregunta.
 - ¿Cuál es la probabilidad de que un alumno conteste correctamente a 10 preguntas?
 - ¿Cuál es la probabilidad de que la primera pregunta correcta sea la cuarta?
 - ¿Cuál es la probabilidad de que la undécima pregunta sea la quinta correcta que contesta un alumno?

Fórmulas:

- Masa de probabilidad **Binomial**: $\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots, n-1, n\}$.
- Masa de probabilidad **Binomial negativa**: $\mathbb{P}(X = x) = \binom{n+x-1}{x} (1-p)^x p^n$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de probabilidad **Poisson**: $\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de prob **Hipergeométrica**: $\mathbb{P}(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$, $x \in \text{Sop}(X) = \{\max(0, n+k-N), \min(k, n)\}$.
- Recta de regresión**: $y = a + bx$, $b = \frac{S_{xy}}{S_x^2}$, $a = \bar{y} - b\bar{x}$; $r = \frac{S_{xy}}{s_x s_y}$; $S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$

	Grado en Ingeniería Informática - Universidade de Santiago de Compostela	
	Asignatura: Estadística	Curso: 2021-2022
	Nombre:	Prueba Ev. Cont. 2
	Apellidos:	Nota:

1. (2.5 puntos) El número de visitas a un determinado portal web ha sido contabilizado en los últimos 25 días. Se sabe que, en cuanto a la variabilidad (medida en desviación típica poblacional), $\sigma = 5$ visitas diarias. ¿Cuál es la probabilidad de que el número medio de visitas al portal en esos 25 días se diferencie de la media real de visitas en menos de 1 visita?

2. Se desea estudiar la media de las duraciones de las conexiones a Internet de un grupo de estudiantes. Para ello, se dispone de una muestra aleatoria de 200 estudiantes, con una duración media de sus conexiones de 4.18 horas/día.


Suponemos que la duración de las conexiones de la población sigue una distribución normal con desviación típica igual a 0.74 horas/día.

a) (3 puntos) Calcula el intervalo de confianza del 95 % para la media poblacional de la duración de las conexiones a Internet.

b) (4.5 puntos) Estudios previos aseguran que la duración media de las conexiones a Internet en el último año fue de 4.07 horas/día. Para una significación del 5 %, ¿constituyen estos resultados una prueba significativa de que la duración media (en horas/día) es mayor que la del último año? ¿Y para una significación del 1 %?

Fórmulas:

- Masa de probabilidad **Binomial**: $\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x \in \text{Sup}(X) = \{0, 1, 2, \dots, n-1, n\}$.
- Masa de probabilidad **Binomial negativa**: $\mathbb{P}(X = x) = \binom{n+x-1}{x} (1-p)^x p^n$, $x \in \text{Sup}(X) = \{0, 1, 2, \dots\}$.
- Masa de probabilidad **Poisson**: $\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \in \text{Sup}(X) = \{0, 1, 2, \dots\}$.
- Masa de prob **Hipergeométrica**: $\mathbb{P}(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$, $x \in \text{Sup}(X) = \{\max(0, n+k-N), \min(k, n)\}$.
- Función de distribución **Gamma**: $1 - \sum_{j=0}^{p-1} \frac{(\lambda x)^j}{j!} e^{-\lambda x}$, $x > 0$, p entero positivo.
- **Recta de regresión**: $y = a + bx$, $b = \frac{S_{xy}}{s_x^2}$, $a = \bar{y} - b\bar{x}$; $r = \frac{S_{xy}}{s_x s_y}$; $S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$.
- Si $X \sim \text{Ber}(p)$: $\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$.
- Si $X \sim N(\mu, \sigma^2)$: $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, $\frac{\bar{X} - \mu}{s/\sqrt{n-1}} \sim t_{n-1}$, $\frac{ns^2}{\sigma^2} \sim \chi_{n-1}^2$.
- Si $X \sim N(\mu_X, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$: $\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{S_T \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$, con $S_T^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$,
 $S_X^2 \sigma_Y^2 / (S_Y^2 \sigma_X^2) \sim F_{n-1, m-1}$.

 Escola Técnica Superior de Enxeñaría	Grado en Ingeniería Informática - Universidade de Santiago de Compostela	
	Asignatura: Estadística	Curso: 2022-2023
	Nombre:	Prueba Ev. Cont. 2
	Apellidos:	Nota:

1. Una empresa de telefonía móvil dota a sus teléfonos de baterías que provienen de dos fábricas distintas, A y B. Las baterías de la fábrica A duran una media de 114 horas, con una desviación típica de 23 horas, mientras que las de la fábrica B tienen una duración media de 109 horas y una desviación típica de 17 horas. Se toma una muestra al azar formada por 63 baterías de la fábrica A y 55 de la fábrica B. Suponiendo la normalidad de los datos:


- ¿cuál es la distribución de la diferencia de medias para la muestra anterior? (1 punto)
- ¿cuál es la probabilidad aproximada de que la diferencia de las medias sea mayor a 18 horas? (2,5 puntos)

2. Una compañía informática asegura que la vida útil de los equipos informáticos que fabrica supera los 4,25 años. Se sabe que la distribución de ese tiempo es normal. Se recogen las vidas útiles de 101 equipos y se observa que el tiempo medio de vida útil fue de 4 años, con desviación típica de 1,25 años.

- Obtén el intervalo de confianza al 95 % para la media de vida útil de los equipos de esa compañía. (2,5 puntos)
- Con un nivel de significación del 5 %, ¿existen pruebas significativas de que la vida útil de los equipos es mayor que 4,25 años? ¿Y con un nivel del 1 %? (4 puntos)

Fórmulas:

- Masa de probabilidad **Binomial**: $\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots, n-1, n\}$.
- Masa de probabilidad **Binomial negativa**: $\mathbb{P}(X = x) = \binom{n+x-1}{x} (1-p)^x p^n$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de probabilidad **Poisson**: $\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de prob **Hipergeométrica**: $\mathbb{P}(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$, $x \in \text{Sop}(X) = \{\max(0, n+k-N), \min(k, n)\}$.
- Función de distribución **Gamma**: $1 - \sum_{j=0}^{p-1} \frac{(\lambda x)^j}{j!} e^{-\lambda x}$, $x > 0$, p entero positivo.
- Recta de regresión**: $y = a + bx$, $b = \frac{S_{xy}}{s_x^2}$, $a = \bar{y} - b\bar{x}$; $r = \frac{S_{xy}}{s_x s_y}$; $S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$.
- Si $X \sim \text{Ber}(p)$: $\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$.
- Si $X \sim N(\mu, \sigma^2)$: $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, $\frac{\bar{X} - \mu}{s/\sqrt{n-1}} \sim t_{n-1}$, $\frac{ns^2}{\sigma^2} \sim \chi_{n-1}^2$.
- Si $X \sim N(\mu_X, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$: $\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{S_T \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$, con $S_T^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$, $S_X^2 \sigma_Y^2 / (S_Y^2 \sigma_X^2) \sim F_{n-1, m-1}$.

 Escola Técnica Superior de Enxeñaría	Grado en Ingeniería Informática - Universidade de Santiago de Compostela	
	Asignatura: Estadística	Curso: 2021-2022
	Nombre:	Examen final
	Apellidos:	Nota:

- Del siguiente conjunto de datos de la variable X , $\{x_i, i = 1, \dots, 6\} = \{2, 3, 4, 4, 5, 6\}$.
 - ¿Podría ser la media menor que 2?
 - Calcula la media aritmética, la moda y la mediana.
 - Calcula la varianza, la desviación típica y el rango muestrales.
 - Calcula la media de los valores transformados $y_i = 3x_i - 2$.
- Unos conocidos productores de café de Colombia utilizan compañías aéreas locales para enviar el café producido desde las montañas al aeropuerto internacional más cercano. Por razones de coste, el 65 % de las veces contratan a la compañía AirWings, mientras que los viajes restantes los realizan con LifeFlight. Ambas compañías poseen aviones Tupolev (la mitad de las aeronaves de AirWings y el 75 % de las de LifeFlight son de este fabricante). Calcula:
 - La probabilidad de que uno de los envíos no se realice en Tupolev.
 - Si el envío desde las montañas ha sido realizado en un Tupolev, calcula la probabilidad de que la compañía que lo ha transportado sea LifeFlight.
 - La probabilidad de que el envío sea con AirWings o en Tupolev.

Fórmulas:

- Masa de probabilidad **Binomial**: $\mathbb{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots, n-1, n\}$.
- Masa de probabilidad **Binomial negativa**: $\mathbb{P}(X = x) = \binom{n+x-1}{x} (1-p)^x p^n$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de probabilidad **Poisson**: $\mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x \in \text{Sop}(X) = \{0, 1, 2, \dots\}$.
- Masa de prob **Hipergeométrica**: $\mathbb{P}(X = x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$, $x \in \text{Sop}(X) = \{\max(0, n+k-N), \min(k, n)\}$.
- Función de distribución **Gamma**: $1 - \sum_{j=0}^{p-1} \frac{(\lambda x)^j}{j!} e^{-\lambda x}$, $x > 0$, p entero positivo.
- Recta de regresión**: $y = a + bx$, $b = \frac{S_{xy}}{S_x^2}$, $a = \bar{y} - b\bar{x}$; $r = \frac{S_{xy}}{s_x s_y}$; $S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$.
- Si $X \sim \text{Ber}(p)$: $\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$.
- Si $X \sim N(\mu, \sigma^2)$: $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, $\frac{\bar{X} - \mu}{s/\sqrt{n-1}} \sim t_{n-1}$, $\frac{ns^2}{\sigma^2} \sim \chi_{n-1}^2$.
- Si $X \sim N(\mu_X, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$: $\frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{S_T \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}$, con $S_T^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$,
 $S_X^2 \sigma_Y^2 / (S_Y^2 \sigma_X^2) \sim F_{n-1, m-1}$.

Examen final Estadística 1ª convocatoria

Grado en Ingeniería Informática – Curso 2020/21

Ejercicio 1. (2 puntos) Se tiene que un conjunto de datos de la variable aleatoria X que mide el número de veces que cada persona utiliza un gel hidroalcohólico al día. El soporte de la variable aleatoria solo tiene 4 posibles valores. Se obtiene una muestra de 16 personas, de entre las cuales sabemos que: 3 personas usan el gel hidroalcohólico 4 veces al día; 2 personas 5 veces al día y 5 personas 6 veces al día. Del resto de personas no tenemos, a priori, más información.

- a) (0.25 puntos) ¿Ante qué tipo de variable estadística estamos? Justifica la respuesta.
- b) (0.75 puntos) ¿Podría ser la media muestral menor o igual que tres? ¿Podría ser la media poblacional menor o igual que tres? Justifica la respuesta.
- c) (0.5 puntos) Si se sabe que el resto de personas, que faltaban por catalogar en la muestra, han respondido que utilizan el gel 2 veces al día. Obtén la tabla de frecuencias (no es necesario incluir las frecuencias acumuladas).
- d) (0.5 puntos) Asumiendo el escenario del apartado c). Proporciona un diagrama de barras. ¿Cuál es la moda?

Ejercicio 2. (3 puntos) Un centro de computación cuenta con 10 nodos de computación independientes, 4 de los cuales son de AMD. Se sabe que la asignación del nodo se hace al azar y que todos tienen, a priori, la misma probabilidad de ser elegidos. Tras enviar dos trabajos al centro de computación, se quiere analizar la variable aleatoria que estudia el número de nodos AMD que son utilizados. Además, se quiere saber cuál es la probabilidad de que se use, únicamente, un nodo de AMD. Para ello se asumen dos escenarios:

-Escenario A: el mismo nodo puede ser usado por más de un trabajo, por tanto, a donde haya ido el primer trabajo no influye en el segundo.

-Escenario B: cada nodo solo puede ser usado por un único trabajo.

- a) (0.75 puntos) Bajo el escenario A, determina el modelo de distribución de la variable de interés y calcula la probabilidad que se pide en el enunciado.
- b) (0.75 puntos) Bajo el escenario B, determina el modelo de distribución de la variable de interés y calcula la probabilidad que se pide en el enunciado.
- c) (0.75 puntos) El 30% de las veces se manda un trabajo a un centro de computación en el escenario A y el 70% a un centro en el escenario B. Sabiendo que únicamente un trabajo de los dos enviados ha acabado en un nodo AMD, ¿cuál es la probabilidad de que dicho trabajo se haya ejecutado en un centro de computación en el escenario A? (Nota: si no fuiste capaz de calcular las probabilidades de los apartados a y b, asume que las conoces)
- d) (0.25 puntos) Se han construido nuevos centros de computación en un nuevo escenario C. ¿Qué condiciones deberían darse sobre estos centros en el escenario C, para que los centros en los escenarios A, B y C formen un conjunto completo de sucesos?
- e) (0.5 puntos) Tras construir estos nuevos centros de computación, no se disponen de nuevos centros en los escenarios A y B. Por tanto, cuando el trabajo no es mandado a los centros en el escenario C, se mandan a un centro en el escenario A el 30% de las veces y el 70 % a un centro en el escenario B. A mayores se sabe que si se manda a un centro en el escenario A, no se manda a un centro en el escenario C. Si el escenario C se escoge con una probabilidad de 0.2, ¿cuál es la probabilidad de que A sea el escenario escogido?

Ejercicio 3. (1.25 puntos) Un servicio de asistencia técnica en carretera ha comprobado que en las mañanas de los fines de semana el número de llamadas que recibe, por término medio, es de 3 llamadas cada hora. Un operario comienza su jornada de sábado a las 8 de la mañana. Suponiendo que las llamadas se realizan de forma independiente y con tas constante:

- ¿Cuál es la probabilidad de que reciba la primera llamada antes de las 8:15? (0.5 puntos)
- Si lleva 10 minutos sin recibir ninguna llamada, ¿cuál es la probabilidad de que reciba una nueva llamada en los siguientes 5 minutos? (0.75 puntos)

Ejercicio 4. (1.25 puntos) Una empresa de material informático está interesada en analizar el tiempo de vida de sus discos duros externos sabiendo que dicho tiempo sigue una distribución Normal de media $\mu = 900$ y desviación típica $\sigma = 95$ días. Se pide:

- Calcula la probabilidad de que el tiempo de vida esté comprendido entre 841 y 916 días. (0.5 puntos)
- ¿Cuál es el tiempo de vida mínimo asociado al 25% de los dispositivos con mayor tiempo de vida? (0.5 puntos)
- Si se tienen 10 discos con tiempos de vida independientes, determina la distribución y los parámetros asociados a la variable tempo total de vida de los 10 dispositivos. (0.25 puntos)

Ejercicio 5. (2.5 puntos) Se realiza un estudio de mercado sobre portátiles convertibles para conocer la opinión de los clientes en función de varias características. Se ha encuestado a 32 usuarios de convertibles con disco duro HDD y 90 con disco SSD y se tuvo en cuenta que dispusiesen o no de un lápiz de escritura de precisión. Se les pidió que valorasen la satisfacción con sus portátiles convertibles de 0 a 20 obteniendo los valores que se resumen en la siguiente tabla:

Calificación	HDD		SSD	
	Media	Cuasi-desv. típica	Media	Cuasi-desv. típica
	12.56	3.14	17.11	4.13
Lápiz escritura	Sí 24	No 8	Sí 71	No 19

A la vista de los resultados y asumiendo que la calificación de los dispositivos sigue una distribución Normal de igual varianza en ambos tipos de discos duros,

- ¿podemos concluir, fijado un nivel de significación del 5%, que existen diferencias significativas entre la puntuación media en convertibles con disco HDD y SSD? Calcula el p-valor y razona la conclusión del contraste. (1 punto)
- Construye un intervalo de confianza de nivel 95% para la proporción poblacional de usuarios de convertible con disco SSD y sin lápiz de escritura de precisión. (0.75 puntos)
- Define los siguientes conceptos relativos a contrastes de hipótesis: error de tipo I, error de tipo II y p-valor. (0.75 puntos)

Examen Estadística - Junio 2010

Grado en Ingeniería Informática

Apellidos: _____ Nombre: _____

Cuestiones

1. **(1 pto)** Se ha recogido información sobre el tiempo, en horas, dedicado a preparar el caso práctico de Estadística, preguntando a 50 alumnos y obteniendo las siguientes medidas características: $Q_1 = 10$, $Q_2 = 12$, $Q_3 = 24$. También se ha registrado un mínimo de $x_{(1)}$ horas y un máximo de 32, sin la presencia de datos atípicos. Indica si son posibles o no las siguientes afirmaciones, justificando la respuesta.
 - a) El 65 % del alumnado ha dedicado menos de 11 horas.
 - b) El 70 % tarda entre 15 y 20 horas.
 - c) El tiempo medio dedicado es superior a 12 horas.
 - d) El rango de variación de los datos es de 13 horas.
2. **(0.5 pto)** Explica qué calculan las siguientes secuencias de comandos de R:
 - a) Secuencia 1:

```
x<-rexp(100,lambda=2)
hist(x,freq=F)
lines(density(x))
abline(v=mean(x))
```
 - b) Secuencia 2:

```
x<-datos$V1;y<-datos$V2
mod<-lm(y~x)
boxplot(mod$residuals)
```
3. **(1.5 ptos)** Sean X e Y dos variables aleatorias con distribuciones $X \sim N(3, 4)$ e $Y \sim N(0, 9)$ y cuya covarianza vale $\sigma_{XY} = 1$. Determina la distribución de $(X - Y)$ y calcula la probabilidad de que X sea mayor que Y .

Ejercicios

4. **(2 ptos.)** Unos conocidos productores de café de Colombia utilizan compañías aéreas locales para enviar el café producido desde las montañas al aeropuerto internacional más cercano. Por razones de coste, el 65 % de las veces contratan a la compañía AirWings, mientras que los viajes restantes los realizan con LifeFlight. Ambas compañías poseen aviones Tupolev (la mitad de las aeronaves de AirWings y el 75 % de las de LifeFlight son de este fabricante). Calcula:
- a) La probabilidad de que uno de los envíos no se realice en Tupolev.
 - b) Si el envío desde las montañas ha sido realizado en un Tupolev, calcula la probabilidad de que la compañía que lo ha transportado sea LifeFlight.
 - c) La probabilidad de que el envío sea con AirWings o en Tupolev.
5. **(2.5 ptos.)** A través de estudios empíricos se ha llegado a la conclusión de que uno de cada cuatro estudiantes de la USC han ido a algún espectáculo cultural durante el año 2010.
- a) Del grupo de los jueves (11 alumnos), ¿cuál es la probabilidad de que 3 hayan asistido a algún espectáculo durante este año?
 - b) De los que normalmente asisten a clase (40 alumnos), ¿cuál es la probabilidad de que 20 hayan asistido?
 - c) ¿Cuál es el número medio de alumnos a los que debemos preguntar antes de encontrar al primero que haya asistido?
6. **(2.5 ptos.)** El número diario de pasajeros extracomunitarios que recibe el aeropuerto de Lavacolla sigue una distribución Normal. Después de hacer un recuento de pasajeros durante dos meses (60 días), se ha obtenido una media de 21 pasajeros y una desviación típica de 20 pasajeros.
- a) Obtén un intervalo de confianza al 95 % para el número medio de pasajeros.
 - b) Si fijamos una significación del 1 %, ¿podemos admitir que el número medio de pasajeros es inferior a 23?

Información:

Publicación de notas: Luns 14/06/10. Campus Virtual e taboleiros da Escola Técnica Superior de Enxeñaría.

Revisión de exámes:

Mércores, 16/06/10 de 16:00 a 18:00. Xoves, 17/06/10 de 10:00 a 12:00.

Examen Estadística - Julio 2010

Grado en Ingeniería Informática

Apellidos: _____ Nombre: _____

Cuestiones

1. **(1 pto.)** Sean A, B dos sucesos definidos sobre el mismo espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$ tales que $\mathbb{P}(A \cup B) = 0'8$, $\mathbb{P}(\bar{A}) = 0'6$, A y B incompatibles.
 - a) $\mathbb{P}(\overline{A \cap B})$
 - b) $\mathbb{P}(\bar{B})$
 - c) ¿Son independientes?
2. **(1 pto.)** Sean X e Y dos variables aleatorias independientes con distribución Normal de media $\mu_X = 2$, $\mu_Y = 3$ y varianzas $\sigma_X^2 = 1$, $\sigma_Y^2 = 2$ respectivamente. Calcula:
 - a) $\mathbb{P}(X > Y)$
 - b) $\mathbb{P}(Y < 3'1)$
 - c) $\mathbb{P}(X > 2 | Y < 3'1)$

Ejercicios

2. **(2 ptos.)** La OMS indica que la *Giardia lamblia* es la causa más común cuando la diarrea es causada por un parásito. En una guardería a la que asisten 84 niños y 56 bebés, todos tuvieron problemas de diarrea por causa de dicho parásito. Se prescribe un medicamento, que cuando se suministra a los niños erradica la diarrea con probabilidad de 0'7 y a los bebés con probabilidad 0'6. Se elige al azar unos de los afectados:
 - a) Probabilidad de erradicar la diarrea.
 - b) Si no se consigue erradicar la diarrea, probabilidad de que el afectado sea un bebé.
 - c) Si se seleccionan 5 afectados. Probabilidad de erradicar en todos ellos la diarrea.

3. **(1.5 ptos.)** El número de fallos de acceso a un servidor de base de datos de Xescampus sigue una distribución de Poisson de media 5 fallos cada 10 minutos. Calcula:
- a) La probabilidad de que en 2 minutos haya al menos 1 fallo.
 - b) La probabilidad de que el tiempo entre dos fallos consecutivos sea inferior a 30 segundos.
 - c) La probabilidad de que haya que esperar más de dos horas para tener 50 fallos.
4. **(1.5 puntos)** El webmaster de SoftFree.com sostiene que más del 70 % de los accesos a su página provienen de búsquedas en Bugle. Tras observar 200 accesos, se ha comprobado que 120 han llegado a través del buscador Bugle. Si fijamos una significación del 10 %, ¿podemos fiarnos del webmaster?
-

Información:

Publicación de notas: Lunes, 19/07/10. Campus Virtual y tableros de la Escuela Técnica Superior de Ingeniería.

Revisión de exámenes: Martes, 20/07/10 a las 12:30 (Dpto. de Estadística e Investigación Operativa, Facultad de Matemáticas).

Exame Estatística - Maio 2011

Grao en Enxeñaría Informática

Apelidos: _____ **Nome:** _____

1. **(1 pto)** Sexan X e Y dúas variables aleatorias definidas sobre $(\Omega, \mathcal{A}, \mathbb{P})$, independentes. Consideremos a nova variable $Z = 2X - 3Y + 6$. Calcula o valor esperado e a varianza de Z supoñendo que X segue unha distribución de Poisson con frecuencia $\lambda = 3$ e Y é unha variable normal de media 3 e varianza 4.
2. **(1.5 ptos)** Sexan A, B sucesos independentes nun mesmo espazo de probabilidade $(\Omega, \mathcal{A}, \mathbb{P})$. Se $\mathbb{P}(\bar{A}) = 0'3$ e $\mathbb{P}(A \cup B) = 0'7$, calcula:
 - a) $\mathbb{P}(B)$
 - b) $\mathbb{P}(A \setminus B)$
 - c) Sexa $C \in (\Omega, \mathcal{A}, \mathbb{P})$. Que condicións deberían darse sobre C para que A e C formen un conxunto completo de sucesos?
3. **(0.5 ptos)** Indica que operacións realizan os seguintes comandos de R:
 - a) `boxplot(rnorm(100))`
 - b) `qnorm(0.5, sd=2)`
 - c) `curve(x, dt(x, df=3), from=-4, to=4)`
 - d) `hist(lm(y~x)$residuals, freq=F)`
4. **(2 ptos)** Unha empresa de reparto de mercadorías cobre diariamente o servizo entre Ferrol e Ribadeo pola N642. Supoñamos que a probabilidade de que sufra un accidente nun día con néboa é 0'02 e nun día sen néboa é 0'004. Certo día dun mes, no que houbo 12 días con néboa e 18 sen néboa, produciuse un accidente. Calcula:
 - a) A probabilidade de que o accidente fora un día con néboa.
 - b) A probabilidade de que o accidente fora un día sen néboa.
5. **(2 ptos)** O Servizo de Atención ao Cliente dunha empresa de telefonía móbil recibe, en promedio, 6 chamadas cada cuarto de hora. Calcula:
 - a) A probabilidade de que nunha hora se reciban máis de 20 chamadas.
 - b) A probabilidade de que haxa que esperar menos de 1 minuto entre dúas chamadas consecutivas.
 - c) O tempo medio de agarda ata recibir 40 chamadas.

6. **(3 pts)** O antivirus AC USB Virus Killer posúe un escáner en tempo real que se activa cada vez que o usuario conecta un novo dispositivo ao ordenador. Nun ordenador onde se instalou o AC USB Virus Killer, probáronse 100 lápises de memoria infectados, eliminando o Recycler en 30 deles.
- a) Calcula un intervalo de confianza ao 95 % para a proporción de lápises de memoria que quedan limpos ao aplicar o AC USB Virus Killer. Que tamaño de mostra teríamos que utilizar para reducir a lonxitude do intervalo á metade da actual, supoñendo que a proporción mostral de lápises desinfectados non varía?
 - b) O fabricante afirma que este antivirus é máis eficaz que os antivirus utilizados habitualmente para eliminar o Recycler. Para comprobar a hipótese, instalouse noutro ordenador un antivirus convencional, probáronse 80 lápises infectados, eliminándose o Recycler en 20. Se fixamos unha significación do 10 %, podemos fiarnos do fabricante?

Información:

Publicación de notas: Martes 14/06/11. Taboleiros da Escola Técnica Superior de Enxeñaría e Campus Virtual.

Revisión de exámes: Luns, 21/06/11 de 10:00 a 12:00 e de 16:00 a 18:00.

Exame Estatística - Xullo 2011

Grao en Enxeñaría Informática

Apelidos: _____ **Nome:** _____

1. **(1 pto)** Tres alumnos do Grao en Enxeñaría Informática quedan os venres pola tarde para xogar aos dardos. Se a probabilidade de acertar no centro da diana para cada un deles é $1/6$, $1/4$ e $1/3$ respectivamente, calcula a probabilidade de que tan só un deles acerte na diana.
2. **(1.5 ptos)** Sexa X unha variable aleatoria Binomial de media 20 e varianza 16. Calcula a probabilidade de que X tome o valor 25.
3. **(1 pto)** Un grupo de manufacturación de turismos encarga un estudo sobre a contaminación acústica dos vehículos que produce nunha cidade onde o 20 % dos coches son pequenos, o 50 % medianos e os restantes son de gama alta. O 40 % dos coches pequenos dispoñen dun silenciador extra, mentres que o 50 % dos medianos e o 60 % dos grandes dispoñen deste silenciador.

Calcula a probabilidade de que un vehículo elixido ao azar teña silenciador extra.

4. **(2 ptos)** O nivel de ruído que provoca un vehículo cun silenciador estándar segue una distribución normal de media 730 decibelios e desviación típica 20 decibelios. No caso de vehículos co silenciador extra, a distribución do nivel de ruído tamén segue unha normal de media 800 decibelios con desviación típica 13 decibelios.

Tendo en conta os datos do exercicio anterior, se eliximos un vehículo ao chou, calcula a probabilidade de que o nivel de ruído sexa inferior a 790 decibelios.

5. **(2 ptos)** A mesma empresa de turismos do exercicio anterior realiza tamén un estudo para comparar o tempo de reacción de frenada dun novo modelo que está fabricando entre condutores de 50 a 55 anos e condutores de 25 a 30 anos. Para o estudo tomaron dúas mostras aleatorias e independentes de 40 persoas de cada grupo, obtendo os seguintes resultados para o tempo de reacción:

- Mostra de 50 a 55 anos: media=4 segundos; varianza=0'189 segundos
- Mostra de 25 a 30 anos: media=3'5 segundos; varianza=0'1785 segundos.

Á vista dos resultados, pódese afirmar que os condutores de 50 a 55 anos teñen un tempo de reacción superior ao dos máis novos?

6. **(1.5 ptos)** As estatísticas de Dropbox indican que se rexistran un promedio de 2 novos usuarios cada 30 segundos. Se supoñemos que este promedio se mantén constante no tempo, calcula:

- a) A probabilidade de que nun minuto se rexistren máis de 3 novos usuarios.
- b) A probabilidade de que, nunha hora, se rexistren máis 235 usuarios.
- c) A probabilidade de ter que esperar máis dun minuto antes de que se rexistre o primeiro usuario.

Información:

Publicación de notas: 18/07/11. Taboleiros da Escola Técnica Superior de Enxeñaría e Campus Virtual.

Revisión de exámes: 20/07/11 de 10:00 a 12:00.

Exame Estatística - Maio 2012

Grao en Enxeñaría Informática

Apelidos: _____ **Nome:** _____

1. **(1 pto)** Sexan X e Y dúas variables aleatorias definidas sobre $(\Omega, \mathcal{A}, \mathbb{P})$, onde $X \sim N(2, 1)$, $Y \sim \text{Pois}(1)$ e o valor esperado do seu produto é $\mathbb{E}(XY) = 2$. Calcula o valor esperado e a varianza da nova variable $W = X - 2Y$. Cales serían o valor esperado e a varianza se o coeficiente de correlación $\rho_{XY} = 1$?
2. **(1.5 pto)** Sexan A , B e C sucesos nun mesmo espazo de probabilidade $(\Omega, \mathcal{A}, \mathbb{P})$ onde $\mathbb{P}(A) = 0'3$, $\mathbb{P}(B) = 0'6$ e $\mathbb{P}(C) = 0'25$. Se A e B son independentes e ademais A e C son incompatibles, calcula:
 - b) $\mathbb{P}(A \cap B|C)$.

Se B e C son tamén incompatibles, calcula:

 - b) $\mathbb{P}(A \cup B \cup C)$.
3. **(0.5 pto)** Indica que operacións realizan os seguintes comandos de R:


```
ns<-500;n<-100
M<-matrix(rexp(ns*n,rate=1),ncol=ns,nrow=n)
v<-apply(M,2,sum)
hist(v,freq=FALSE);lines(density(v),col=2)
```
4. **(2 pto)** Nunha análise de efectividade de programas antivirus para o troiano Bifrost, o 57 % dos usuarios utiliza o antivirus tipo A (cunha efectividade do 99 %), o 38 % teñen o antivirus tipo B (efectividade do 95 %) e o resto dos usuarios, teñen outros antivirus cunha efectividade do 93 %. Se eliximos un usuario ao chou e comprobamos que o antivirus foi efectivo, calcula a probabilidade de que ese usuario non teña o antivirus tipo A.
5. **(1.5 pto)** O portal de autenticación da AEAT a través do que se xestionan os borradores da declaración da renda, recibe durante os meses de maio e xuño unha media de 6 peticións por minuto. Calcula:
 - a) A probabilidade de que en medio minuto se reciban menos de 2 peticións.
 - b) A probabilidade de que nunha hora se reciban máis de 300 peticións.

6. **(3.5 ptos)** Da enquisa realizada polo alumnado de GrEI sobre se as/os galegas/os querían un adianto das eleccións autonómicas, no sector de idade ata 23 anos (75 persoas), obtívose que un 49'3 % serían favorables a un adianto electoral.

- a) Calcula un intervalo de confianza ao 90 % para a proporción mozas/mozos de ata 23 anos que queren un adianto electoral.

Por outra banda, no grupo de maiores de 23 anos (76 persoas), eran partidarios do adianto electoral o 43'6 %.

- b) Pódese afirmar, cunha significación do 5 %, que a porcentaxe de partidarios do adianto electoral nos mozos é maior que no grupo de maiores de 23 anos?

Información:

Publicación de notas: Martes 12/07/11. Taboleiros da Escola Técnica Superior de Enxeñaría e Campus Virtual.

Revisión de exámes: Xoves, 14/07/11 de 10:00 a 12:00 e de 16:00 a 18:00. Dpto. de Estatística e Investigación Operativa. Facultade de Matemáticas.

Examen Estadística - Julio 2012

Grado en Ingeniería Informática

Apellidos: _____ **Nombre:** _____

1. **(1.25 pto)** Sean A , B y C sucesos en un mismo espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$ donde $\mathbb{P}(A) = 0'3$, $\mathbb{P}(B) = 0'6$ y $\mathbb{P}(C) = 0'25$. Si A y C son independientes, A y B son incompatibles y $\mathbb{P}(B \cup C) = 0'8$ calcula $\mathbb{P}(A \cup B|C)$.
2. **(1.25 pto)** Determina la probabilidad de realizar una prueba con éxito si se sabe que al repetirla 24 veces, la probabilidad de obtener 4 y 5 éxitos es la misma.
3. **(0.5 pto)** Indica qué operaciones realizan los siguientes comandos de R:

```
y<-rexp(500,rate=2)
hist(y,freq=F)
s<-seq(0,10,length=250)
lines(dexp(s,rate=2)~s,col=2)
```
4. **(2 pto)** En el período estival, las touroperadoras VacaSol, SolPlaya y TourSol fletan vuelos charter desde Lavacolla, copando respectivamente el 25 %, 60 % y 15 % de los turistas. Aunque su coste es inferior al de los vuelos en líneas regulares, los charter de las touroperadoras suelen tener horarios intempestivos y sufrir retrasos. En concreto, el año pasado, un 30 % de los vuelos de VacaSol salieron con retraso, al igual que el 70 % de los de SolPlaya y el 45 % de los de TourSol.
 - a) Si Pepe decide organizar sus vacaciones con una de estas touroperadoras, ¿cuál es la probabilidad de que su vuelo salga con retraso?
 - b) Una vez que Pepe ha llegado a destino, coincide en el hotel con Paco que también ha viajado desde Santiago, de manera independiente. Sabiendo que ambos han viajado en un vuelo con retraso, ¿cuál es la probabilidad de que Pepe haya viajado con TourSol y Paco lo haya hecho con SolPlaya?
5. **(2.5 pto)** La aplicación de solicitud de cita previa para renovación del DNI soporta una media de 4 peticiones cada 10 minutos. Calcula:
 - a) La probabilidad de que en dos minutos se reciba alguna petición.
 - b) La probabilidad de que haya que esperar más de dos minutos entre dos peticiones consecutivas.
 - c) Si debido al procedimiento de autenticación empleado, la aplicación soporta un máximo de 40 peticiones cada media hora, ¿cuál es la probabilidad de que el servicio falle?

6. **(2.5 ptos)** En una encuesta realizada a 50 alumnos/as de primer curso de GrEI, se deduce que han dedicado al estudio una media de 10'5 horas diarias durante el mes de junio, con una cuasidesviación de 3'25 horas.

- a) Suponiendo que el tiempo dedicado al estudio se distribuye según una normal, calcula un intervalo de confianza al 99 % para el tiempo medio diario de horas dedicadas al estudio.

Por otro lado, en una encuesta realizada a 100 alumnos/as de otras titulaciones, se obtiene que el tiempo medio dedicado es de 9 horas diarias, con una cuasidesviación de 3'5 horas. Suponiendo que las desviaciones sobre el tiempo medio son las mismas en ambos grupos:

- b) ¿Se puede afirmar, con una significación del 5 %, que el tiempo medio diario dedicado al estudio es mayor en el alumnado de GrEI que en otras titulaciones?

Información:

Publicación de notas: Lunes 09/07/12. Tablones informativos de la Escuela Técnica Superior de Ingeniería y Campus Virtual.

Revisión de exámenes: Miércoles, 11/07/12 de 10:00 a 12:00 y de 16:00 a 18:00. Dpto. de Estadística e Investigación Operativa. Facultad de Matemáticas.

Apellidos y nombre:

1. (1 punto) La duración de las llamadas a un servicio de atención al cliente recogidas a lo largo de 30 días varía entre un mínimo de 10 minutos y un máximo de 60 minutos, con una media de 30 minutos. En dicha muestra, se conoce que el rango intercuartílico es de 25 minutos y tras un análisis descriptivo, se ha determinado que no existen llamadas de duración atípica. Con esta información, determina el valor del primer y el tercer cuartil e interpreta los resultados. ¿Podrías dibujar un diagrama de caja con los datos de los que dispones?
2. (1 punto) Sean A y B dos sucesos independientes en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbb{P})$. Si $\mathbb{P}(A) = 1/3$, $\mathbb{P}(B) = 2/4$ y la de la intersección $3/8$. Calcula:
 - a) La probabilidad de que ocurra alguno de los dos sucesos.
 - b) La probabilidad de que no ocurra ni A ni B .
 - c) La probabilidad de que no ocurra A o bien no ocurra B .
3. (1.75 puntos) En el mercado de los coches de alquiler, un quinto del mercado es copado por la compañía A, mientras que el resto de la cuota se lo reparten a partes iguales las compañías B y C. Además, el 20 % de los coches alquilados por A son eléctricos, mientras que en las compañías B y C dicho porcentaje desciende hasta el 5%.
 - a) Calcula la probabilidad de que un coche alquilado no sea eléctrico.
 - b) ¿Cuál es la probabilidad de que un coche eléctrico provenga de la compañía A?
 - c) ¿Cuál es la probabilidad de que un coche eléctrico provenga de la compañía A o de la compañía B?
4. (1.5 puntos) En una fábrica de refrescos el volumen que una máquina de llenado automático deposita en latas tiene una distribución normal con media 34 cl. y desviación típica 1.5 cl.
 - a) Si se desechan aquellas latas que tienen menos de 33 cl., ¿cuál es la proporción de latas desechadas?
 - b) Si tomamos 10 latas llenadas con la máquina ¿cuál es la probabilidad de que ninguna sea desechada?
 - c) Si ahora tomamos 500 latas llenadas con la máquina ¿cuál es la probabilidad de que al menos 100 sean desechadas?

Indicación: En este ejercicio puedes utilizar comandos de R para responder. En el caso de utilizar los comandos relacionados con la normal, sólo se podrá hacer uso de la normal estándar (media 0 y varianza 1).

5. (1.75 puntos) Una encuesta previa a las elecciones, realizada sobre 100 votantes, daba al partido Independientes por Altamira (IpA) un 52 % de los votos en su ayuntamiento. Sin embargo, una vez realizado el recuento de votos, se determinó que el porcentaje real efectivo fue de diez puntos porcentuales menos, con un 42 %. Con una significación del 5 %, ¿podríamos decir que la encuesta acertó el resultado?

Con los datos anteriores, ¿podrías dar un intervalo de confianza para la proporción de votantes de IpA al 90 %? Interpreta el resultado obtenido.

Indicación: *En este ejercicio puedes utilizar comandos de R para responder. En el caso de utilizar los comandos relacionados con la normal, sólo se podrá hacer uso de la normal estándar (media 0 y varianza 1).*

Nota: Las notas se publicarán el 05/06/2015, previsiblemente por la tarde. Se pondrá el documento con las cualificaciones en el campus virtual y en los tablones de la ETSE. Para la revisión de exámenes habrá dos turnos. El primero será el día 08/06/2015 de 10:00 a 11:00 y el segundo el 09/06/2015, de 9:00 a 11:00. La revisión será en el Dpto. de Estadística e Investigación Operativa (Facultad de Matemáticas).
