Versión A

Respuestas Ejercicio 1 Verdadero o Falso

1	2	3	4	5	6	7	8	9	10
F	F	F	V	F	F	F	F	V	F

Respuestas Ejercicios Multiple Opción (Ejercicios 2 al 11)

E. 2	E. 3	E. 4	E. 5	E. 6	E. 7	E. 8	E. 9	E. 10	E. 11
B	B	C	A	C	B	C	A	C	D

Versión B

Respuestas Ejercicio 1 Verdadero o Falso

1	2	3	4	5	6	7	8	9	10
V	F	F	F	V	F	F	F	F	F

Respuestas Ejercicios Multiple Opción (Ejercicios 2 al 11)

E. 2	E. 3	E. 4	E. 5	E. 6	E. 7	E. 8	E. 9	E. 10	E. 11
B	C	C	A	C	B	B	A	A	D

Ejercicio 1 (Verdadero o falso: 20 puntos)

1. Una combinación lineal de dos funciones de densidad es una función de densidad.
2. Si $X \leq Y$ entonces $Var(X) \leq Var(Y)$
3. Si $\mathbb{E}(X) \leq \mathbb{E}(Y)$ entonces $X \leq Y$
4. $\overline{X}_n$ es un estimador insesgado de $\mathbb{E}(X_1)$
5. Si X es continua entonces $F_X(0) = \frac{1}{2}$
6. Si B_1 y B_2 son eventos independientes entonces $\mathbb{P}(A|B_1 \cap B_2) = \mathbb{P}(A|B_1)\mathbb{P}(A|B_2)$
7. Si $Var(X^2) = E(X^2)$ entonces $X = 0$.
8. Para todo suceso A se tiene que $\mathbb{P}(A|A^c) = \mathbb{P}(A)$
9. $\overline{X_{n+1}} = \overline{X_n}\frac{n}{n+1} + \frac{1}{n+1}X_{n+1}$
10. El sesgo de un estimador siempre es positivo o nulo.

Ejercicio 2 (8 puntos)

Sea X una variable aleatoria con distribución normal estándar ($\mathcal{N}(0,1)$) y sea Y una variable aleatoria, independiente de X que toma sólo los valores 1 y -1 con igual probabilidad. A partir de estas dos variables definimos $Z = X.Y$. Entonces:

(**A**) Z es normal estándar, $Cov(X,Z) = 0$ y Z es independiente de X.

(**B**) Z es normal estándar, $Cov(X,Z) = 0$ y Z no es independiente de X.

(**C**) Z no es normal estándar, $Cov(X,Z) = 0$ y Z es independiente de X.

(**D**) Z no es normal estándar, $Cov(X,Z) \neq 0$ y Z es independiente de X.

Solución:

$$Cov(X,Z) = E(XZ) - E(X)E(Z) = E(X^2Y) - E(X)E(XY) = E(X^2)E(Y) - E^2(X)E(Y) = 0 - 0 = 0.$$

Por otra parte, para $t < 0$ se tiene

$$P\{Z \leq t\} = P\{XY \leq t\} = P\{X \leq t, Y = 1\} + P\{X \geq -t, Y = -1\} =$$
$$0,5.P\{X \leq t\} + 0.5P\{X \geq -t\} = P\{X \leq t\} = \Phi(t)$$

de forma análoga se ve que, si $t > 0$ se cumple

$$P\{Z \leq t\} = P\{X \leq t\} = \Phi(t)$$

Por lo tanto Z es normal estándar.

Sin embargo X y Z no son independientes:

$$P\{|X| < 1, |Z| > 1\} = 0 \neq P\{|X| < 1\}P\{|Z| > 1\}$$

Ejercicio 3 (8 puntos)

Una moneda tiene probabilidad de cara igual a θ, un parámetro desconocido. Se desea hacer el siguiente test de hipótesis sobre el valor de θ:

$$\begin{cases} H_0 : \theta = 0.3 \\ H_A : \theta > 0.3 \end{cases}$$

Suponga que se lanza la moneda 5 veces y se obtienen 3 caras. El p-valor es igual a:

(**A**) 0.324 (**B**) 0.162 (**C**) 0.132 (**D**) 0.030

Solución: Llamemos X al número de caras en 5 lanzamientos de la moneda. Como es un test a una sola cola, el p-valor es $P_{H_0}(X \geq X_{obs}) = P_{H_0}(X \geq 3) = 0.132 + 0.028 + 0.002 = 0.162$

Ejercicio 4 (8 puntos)

Se considera una población numerosa donde se quiere estimar p, la probabilidad de que una persona presente un consumo de alcohol problemático. Se sabe que aquellos que lo presentan lo niegan si se les pregunta directamente.

Para estimar p se ideó el siguiente procedimiento: de una urna con 5 tarjetas, 2 contienen la pregunta (A)¿usted presenta un consumo de alcohol problemático? y las 3 restantes contienen la pregunta (B) ¿usted es moderado en el consumo de alcohol?.

Luego se toma al azar una muestra de n personas. Suponiendo que cada una es capaz de darse cuenta y asumir en caso de tener un problema de consumo de alcohol. A cada una se le pide que saque una tarjeta de la urna, lea la pregunta y conteste “Sí” o “No”, y que luego devuelva la tarjeta a la urna (de esta forma la urna se mantiene siempre con 5 tarjetas). La tarjeta sólo es vista por quien contesta, es decir, el encuestador sólo escucha un “Sí” ó un “No”, con lo cual la persona que responde mantiene su privacidad intacta.
Se supone que todos los encuestados siguen rigurosamente las instrucciones.

Suponiendo que de una muestra de 1000 personas, 580 responden “Sí”, un estimador de p es:

(**A**) 0.58 (**B**) 0.42 (**C**) 0.10 (**D**) 0.12

Solución: Hay dos formas de que la persona encuestada responda “Sí”: que tenga problemas de consumo de alcohol y que le toque una tarjeta con la pregunta A (eso tiene probabilidad $(2/5).p$), la segunda forma es que no tenga problemas de consumo de alcohol y que le toque una tarjeta con la pregunta B (eso tiene probabilidad $(3/5).(1-p)$). De modo que la probabilidad de que una persona responda afirmativamente a la pregunta que le toque es, sumando las probabilidades de las dos formas $3/5-p/5$. El estimador de la probabilidad de una respuesta afirmativa es 580/1000=0.58. Si igualamos este número a $3/5-\hat{p}/5$, donde $\hat{p}$ es el estimador de p, obtenemos $\hat{p}=0.10$.

Ejercicio 5 (8 puntos)

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria exponencial de parámetro α. Se define la distribución empírica de la muestra

$$F_n(x) = \frac{\#\{X_i \leq x\}}{n},$$

Entonces $\mathbb{E}(F_n(x)) =$

(**A**) $1-e^{-\alpha x}$ (**B**) $n(1-e^{-\alpha x})$ (**C**) $e^{-\alpha x}$ (**D**) $ne^{-\alpha x}$

Solución: Usando la definición de la distribución empírica se tiene que $nF_n(x) = \sum_{i=1}^n B_i(x)$, donde las $B_i(x)$ son variables aleatorias de Bernoulli independientes de parámetro $F(x) = (1-e^{-\alpha x})$. Entonces $F_n(x) = \frac{1}{n}\sum_{i=1}^n B_i(x)$. Por la linealidad de la esperanza y la esperanza de una variable aleatoria de Bernoulli, se tiene que $\mathbb{E}(F_n(x)) = F(x) = 1-e^{-\alpha x}$.

Ejercicio 6 (8 puntos)

En una planta de elaboración de quesos se toma para control de calidad un queso de la línea de producción cada 30 minutos. Se considera que el queso no pasó el control si pesa más de 8.08 kg o menos de 7.86 kg, y este es rechazado.

Asumiendo que los pesos de los quesos siguen una distribución normal de esperanza 8 kg y desvío estándar 50 gramos, la probabilidad de no rechazar ningún queso durante una jornada de 8 horas es:

(A) 0.057 **(B)** 0.156 **(C)** 0.388 **(D)** 0.612

Solución:

$$\begin{aligned}
\mathbb{P}(\textit{rechazar un queso}) &= \mathbb{P}(X > 8.08) + \mathbb{P}(X < 7.86) \\
&= \mathbb{P}\left(\frac{X-8}{0.05} > \frac{8.08-8}{0.05}\right) + \mathbb{P}\left(\frac{X-8}{0.05} < \frac{7.86-8}{0.05}\right) \\
&= 1-\phi(1.6)+\phi(-2.8) = 1-\phi(1.6)+1-\phi(2.8) \\
&= 2-0.9452-0.9974 = 0.0574
\end{aligned}$$

Luego la probabilidad de no rechazar ningún queso en 8 horas si se toman muestras cada 30 minutos es la probabilidad de no rechazar quesos en 16 muestras.

$$\mathbb{P}(\textit{no rechazar ningn queso}) = (1-0.0574)^{16} \approx 0.388$$

Ejercicio 7 (8 puntos)

Sean X e Y dos variables aleatorias tales que :

- (X,Y) es normal bivariado
- $X \sim \mathcal{N}(0,1)$
- $Y \sim \mathcal{N}(0,1)$
- La correlación entre X e Y vale $\rho = \frac{1}{4}$.

La probabilidad de que X sea mayor que el doble de Y dado que X es mayor que 0 es:

(A) 0.42 **(B)** 0.58 **(C)** 0.34 **(D)** 0.66

Solución: Se debe calcular $P(X > 2Y|X > 0) = \frac{P(X>2Y,X>0)}{P(X>0)}$.

- $\mathbb{P}(X > 0) = \frac{1}{2}$
- Como $Y = \rho X + \sqrt{1-\rho^2}Z$ con X y Z independientes, el numerador se puede escribir como: $\mathbb{P}\left(X(1-2\rho) > 2\sqrt{1-\rho^2}Z, X > 0\right) = \mathbb{P}\left(Z < \frac{1-2\rho}{\sqrt{1-\rho^2}}X, X > 0\right)$. En este paso se gana que ahora X y Z son independientes, por lo que la normal bi-variada (X,Z) tiene simetría rotacional.

  Teniendo esto en cuenta, se puede graficar la región de interés (asociada al éxito de la probabilidad en el numerador) según se ve en la figura 1.

  Como $\theta_1 = arctg\left(\frac{1-2\rho}{2\sqrt{1-\rho^2}}\right) = 14^{\text{o}}$, la probabilidad del numerador vale: $\frac{\theta_1+90^{\text{o}}}{360^{\text{o}}} = 0.288$.

  Por lo tanto, la probabilidad de la pregunta vale 0.58.

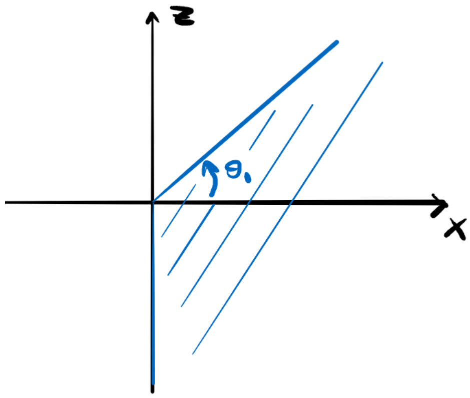


Figura 1: Diagrama de región de interés en el dominio (X,Z).

Ejercicio 8 (8 puntos)

Los siguientes datos correponden a una muestra aleatoria simple de una variable aleatoria Y de la que se sabe $\mathbb{E}(Y) = \frac{\alpha}{\beta}$ y $Var(Y) = \frac{\alpha}{\beta^2}$:

7	14	21	28	35

Una estimación de α por el método de los momentos es:

(A) 3.9 **(B)** 4.2 **(C)** 4.5 **(D)** 4.8

Solución: Planteamos el sistema de ecuaciones igualando los momentos muestrales y poblacionales:

$$\overline{Y} = 21 = \mathbb{E}(Y) = \frac{\alpha}{\beta}$$

$$\frac{1}{n}\sum_i Y_i^2 = 539 = Var(Y) + \mathbb{E}(Y)^2 = \frac{\alpha}{\beta^2} + \frac{\alpha^2}{\beta^2}$$

Despejando en el sistema obtenemos $\alpha = 4.5$.

Ejercicio 9 (8 puntos)

Sean dos variables aleatorias independientes X, con distribución uniforme en $[0,a]$ e Y, con distribución uniforme en $[0,b]$ tal que $0 < b < 1 < a$. Entonces la probabilidad $\mathbb{P}\left(\frac{X}{X+Y} \leq 1/3\right) =$.

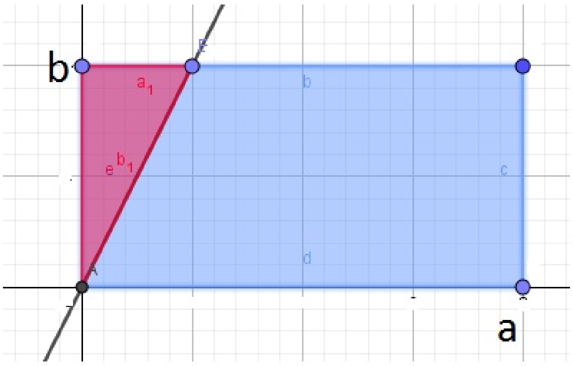
(A) $\dfrac{b}{4a}$ **(B)** $\dfrac{a}{4b}$ **(C)** $\dfrac{b}{3a}$ **(D)** $\dfrac{a}{3b}$

Solución:

$$\mathbb{P}\{\frac{X}{X+Y} \leq 1/3\} = \mathbb{P}\{X \leq \frac{1}{3}X + \frac{1}{3}Y\} = \mathbb{P}\{(1-\frac{1}{3})X \leq \frac{1}{3}Y\} = \mathbb{P}\{2X \leq Y\}$$

Esta probabilidad corresponden al área relativa del triángulo que determinan el semiplano $X \leq Y$ y el rectángulo soporte de la distribución del vector (X,Y). El área del triángulo vale $b^2/4$, al dividirla por el área del rectángulo ab obtenemos

$$\mathbb{P}\{\frac{X}{X+Y} \leq 1/3\} = \mathbb{P}\{2X \leq Y\} = \frac{b}{4a}$$



Ejercicio 10 (8 puntos)

Un dado balanceado que tiene una cantidad k desconocida de caras. Se arroja el dado 8 veces (cada vez independiente entre si), obteniendo los resultados:

3	1	2	3	3	4	7	4

Una estimación de k por el método de máxima verosimilitud es:

(A) 3 **(B)** 4 **(C)** 7 **(D)** 8

Solución: Para un dado balanceado, la fpp es:

$$p(x;k) = \begin{cases} \frac{1}{k}, & \text{si } x = 1..k \\ 0, & \text{sino} \end{cases} \tag{1}$$

Como $L(n) = \Pi_{i=1}^{n} p(x_i;k)$, en primer lugar k debe ser mayor que $max\{x_1, x_2, ..., x_n\}$ para que $L(k) > 0$. Por otro lado, en este caso: $L(n) = \left(\frac{1}{k}\right)^n = \frac{1}{k^n}$. Esta función es decreciente según k, con lo cual $\hat{k} = argmax_k\{L(k)\} = min_k\{L(k) > 0\} = max\{x_1, x_2, ..., x_n\} = 7$.

Ejercicio 11 (8 puntos)

La siguiente tabla muestra el resumen de cinco números de las notas de un examen de inglés para una clase con 16 estudiantes

(R):

min	q_i	m	q_s	max
39	45	49	55	59

Considere los tres conjuntos de datos representados por diagramas de tallo y hojas que se muestran a continuación (tallo = decenas, hojas = unidades)

(D_1):

3	9
4	4
4	557799
5	234
5	55689

(D_2):

3	9
4	44
4	5779
5	023334
5	569

(D_3):

3	99
4	4
4	57799
5	233
5	56789

Indicar cuáles de estos conjuntos de datos se corresponden con el resumen numérico (R).

(A) Sólo (D_1)

(B) Sólo (D_3)

(C) Sólo (D_2) y (D_3)

(D) Sólo (D_1) y (D_3)

Solución: De las definiciones de cuantiles vemos que los resúmenes numéricos de los tres conjuntos de datos son:

(D_1):

min	q_i	m	q_s	max
39	45	49	55	59

(D_2):

min	q_i	m	q_s	max
39	45	50	53	59

(D_3):

min	q_i	m	q_s	max
39	45	49	55	59

Por lo que solamente (D_1) y (D_3) se corresponden al resumen numérico.

Tabla de $\Phi(z)$ (normal estándar)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9958	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Tabla de χ^2

	Probabilidad de cola derecha $P(\chi^2 \geq c)$										
GdL	**.25**	**.20**	**.15**	**.10**	**.05**	**.025**	**.02**	**.01**	**.005**	**.0025**	**.001**
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.52
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59

Tabla de t Student

	Probabilidad de cola derecha $P(t \geq c)$										
GdL	**.25**	**.20**	**.15**	**.10**	**.05**	**.025**	**.02**	**.01**	**.005**	**.0025**	**.001**
1	1.00	1.38	1.96	3.08	6.31	12.71	15.89	31.82	63.66	127.32	318.31
2	0.82	1.06	1.39	1.89	2.92	4.30	4.85	6.96	9.92	14.09	22.33
3	0.76	0.98	1.25	1.64	2.35	3.18	3.48	4.54	5.84	7.45	10.21
4	0.74	0.94	1.19	1.53	2.13	2.78	3.00	3.75	4.60	5.60	7.17
5	0.73	0.92	1.16	1.48	2.02	2.57	2.76	3.36	4.03	4.77	5.89
6	0.72	0.91	1.13	1.44	1.94	2.45	2.61	3.14	3.71	4.32	5.21
7	0.71	0.90	1.12	1.41	1.89	2.36	2.52	3.00	3.50	4.03	4.79
8	0.71	0.89	1.11	1.40	1.86	2.31	2.45	2.90	3.36	3.83	4.50
9	0.70	0.88	1.10	1.38	1.83	2.26	2.40	2.82	3.25	3.69	4.30
10	0.70	0.88	1.09	1.37	1.81	2.23	2.36	2.76	3.17	3.58	4.14

FPP de la binomial $\mathrm{Bin}(5, \theta)$ para varios valores de θ.

x	0	1	2	3	4	5
$\theta = .1$	.590	.328	.073	.008	.000	.000
$\theta = .2$	.328	.410	.205	.051	.006	.000
$\theta = .3$	.168	.360	.309	.132	.028	.002
$\theta = .4$	.078	.259	.346	.230	.077	.010
$\theta = .5$	.031	.156	.313	.313	.156	.031
$\theta = .6$	.010	.077	.230	.346	.259	.078
$\theta = .7$	.002	.028	.132	.309	.360	.168
$\theta = .8$	.000	.006	.051	.205	.410	.328
$\theta = .9$	.000	.000	.008	.073	.328	.590