Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2011

Curso	**:**	**Probabilidad y Estadística**
Sigla	**:**	**EYP1113**
Pauta	**:**	**I3**
Profesor	**:**	**Ricardo Aravena (Sec. 1 y 3), Lorena Correa (Sec. 4) y Ricardo Olea (Sec. 2)**
Ayudantes	**:**	**Tamara Fernandez Aguilar, Felipe Fuentes Astudillo y Claudia Reyes Vizcarra.**

Problema 1

Una empresa realiza un test de 80 preguntas (con dos alternativas, una de las cuales es correcta) a dos aspirantes a una jefatura que se necesita cubrir urgentemente. Por algún motivo se filtraron 44 de las preguntas con sus respectivas respuestas y llegaron a manos de un empleado de menor jerarquía. Este empleado muy amigo de uno de los evaluados (A) le entrega las 44 preguntas, pero por si acaso, entrega 16 de estas preguntas al otro evaluado (B). Suponga que ni A y B supieron del trato del empleado con su competencia. Considere que A y B contestan las preguntas que conocían correctamente con probabilidad 1 y las que no conocían previamente con probabilidad 1/2. Si las preguntas son totalmente independientes:

(a) **[2.0 Ptos.]** Calcule aproximadamente la probabilidad que B conteste correctamente por lo menos 50 preguntas.

(b) **[2.0 Ptos.]** Encuentre la distribución de la suma de las respuestas correctas no conocidas previamente por A y B. Identifique sus parámetros.

(c) **[2.0 Ptos.]** Calcule aproximadamente la probabilidad que el número de preguntas que no conocían previamente y que fueron contestadas correctamente por B y A en total suman más de 40 respuestas correctas.

Solución

(a) Sea X_B la variable aleatoria que denota el número de preguntas que B contesta correctamente entre las 64 preguntas que no conocía previamente. Por lo tanto, dada la independencia entre preguntas, se tiene que

$$X_B \sim \text{Binomial}(n = 64, p = 1/2) \quad \textbf{[0.5 Ptos.]}$$

Por otra parte, una binomial puede ser generada por la suma de n variables aleatorias i.i.d Bernoulli(p). Para nuestro caso

$$X_B = \sum_{i=1}^{n} X_i, \quad \text{donde } X_i \overset{\text{iid}}{\sim} \text{Bernoulli}(p = 1/2). \quad \textbf{[0.2 Ptos.]}$$

Dado que la probabilidad de responder correctamente las 16 preguntas que conocía es uno, la probabilidad solicitada corresponde entonces a

$$\begin{aligned} P\left(X_B \geq 34\right) &= P\left(\frac{X_B - n \cdot p}{\sqrt{n} \cdot \sqrt{p(1-p)}} > \frac{34 - n \cdot p}{\sqrt{n} \cdot \sqrt{p(1-p)}}\right) \quad \textbf{[0.2 Ptos.]} \\ &= P\left(\frac{X_B - 64 \cdot 0{,}5}{\sqrt{64} \cdot \sqrt{0{,}5(1-0{,}5)}} \geq \frac{34 - 64 \cdot 0{,}5}{\sqrt{64} \cdot \sqrt{0{,}5(1-0{,}5)}}\right) \quad \textbf{[0.2 Ptos.]} \end{aligned}$$

$$
\begin{aligned}
P\left(X_B \geq 34\right) &= P\left(\frac{X_B - 32}{\sqrt{16}} \geq \frac{34-32}{\sqrt{16}}\right) \quad \textbf{[0.2 Ptos.]}\\
&= 1 - P\left(\frac{X_B - 32}{\sqrt{16}} < \frac{34-32}{\sqrt{16}}\right) \quad \textbf{[0.2 Ptos.]}\\
&\approx 1 - \Phi\left(\frac{34-32}{\sqrt{16}}\right), \quad \text{por teorema central de límite} \quad \textbf{[0.2 Ptos.]}\\
&= 1 - \Phi(0{,}5) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - 0{,}6915 \quad \textbf{[0.1 Ptos.]}\\
&= 0{,}3085 \quad \textbf{[0.1 Ptos.]}
\end{aligned}
$$

(b) Sea X_A la variable aleatoria que denota el número de preguntas contestadas correctamente entre las 36 preguntas que A no conocía previamente. Por lo tanto se tiene que

$$X_A \sim \text{Binomial}(n = 36,\, p = 1/2). \quad \textbf{[0.5 Ptos.]}$$

Sea la variable aleatoria $Y = X_A + X_B$, la cual tiene distribución de probabilidad

$$\text{Binomial}(n = 100,\, p = 1/2) \quad \textbf{[0.5 Ptos.]}$$

por ser suma de variables aleatorias independientes Bernoulli(p).

Demostración por inducción **[1.0 Ptos.]**

i. Supongamos que $X \sim \text{Binomial}(n,\, p)$ e $Y \sim \text{Binomial}(m = 1,\, p)$, entonces

$$
\begin{aligned}
P(Z = z) &= \sum_{y=0}^{1} P(X = z - y,\, Y = y)\\
&= \sum_{y=0}^{1} P(X = z - y)\, P(Y = y)\\
&= \binom{n}{z-0} p^{z-0}\,(1-p)^0 + \binom{n}{z-1} p^{z-1}\,(1-p)^1\\
&= p^z\,(1-p)^{n-z+1}\left[\binom{n}{z} + \binom{n}{z-1}\right]\\
&= \binom{n}{z} p^z\,(1-p)^{n-z+1}\left[1 + \frac{z}{n-z+1}\right]\\
&= \binom{n}{z} p^z\,(1-p)^{n-z+1}\,\frac{n+1}{n-z+1}\\
&= \frac{n!}{z!\,(n-z)!}\,p^z\,(1-p)^{n-z+1}\,\frac{n+1}{n-z+1}\\
&= \frac{(n+1)!}{z!\,(n-z+1)!}\,p^z\,(1-p)^{n-z+1}\\
&= \binom{n+1}{z} p^z\,(1-p)^{n-z+1}
\end{aligned}
$$

Es decir

$$Z \sim \text{Binomial}(n+1,\, p)$$

ii. Supongamos que para $m = k$, se cumple que $Z \sim \text{Binomial}(n+k,\, p)$ y probar que para $m = k+1$, $Z \sim \text{Binomial}(n+k+1,\, p)$.

(c) Se pide $P\left(X_B + X_A > 40\right) = P\left(Y > 40\right)$. **[0.5 Ptos.]**

Nuevamente por teorema central del límite se tiene que

$$\begin{aligned}
P(Y > 40) &= P\left(\frac{Y - n\cdot p}{\sqrt{n}\cdot\sqrt{p(1-p)}} > \frac{40 - n\cdot p}{\sqrt{n}\cdot\sqrt{p(1-p)}}\right)\\
&= P\left(\frac{Y - 100\cdot 0{,}5}{\sqrt{100}\cdot\sqrt{0{,}5(1-0{,}5)}} > \frac{40 - 100\cdot 0{,}5}{\sqrt{100}\cdot\sqrt{0{,}5(1-0{,}5)}}\right) \quad \textbf{[0.5 Ptos.]}\\
&\approx P\left(Z > -2\right), \quad \text{con } Z \sim \text{Normal}(0,\,1) \quad \textbf{[0.5 Ptos.]}\\
&= 1 - P(Z \leq -2)\\
&= 1 - [1 - \Phi(2)]\\
&= \Phi(2)\\
&= 0{,}9772 \quad \textbf{[0.5 Ptos.]}
\end{aligned}$$

+ 1 Punto Base

Problema 2

El voltaje de un tipo de circuito se comporta como una variable aleatoria uniforme en el intervalo $[0, \theta]$. Sean $x_1, \ldots, x_n$ realizaciones del voltaje X que presentaron n circuitos de manera independiente.

(a) **[1.0 Ptos.]** Determine el estimador de momentos para el parámetro θ.

(b) **[2.0 Ptos.]** Obtenga el valor esperado y varianza del estimador obtenido en (a).

(c) **[2.0 Ptos.]** Para la distribución propuesta en el enunciado, el estimador máximo verosímil de θ, corresponde al $\max\{x_1, \ldots, x_n\}$. Determine la distribución de este estimador (identifique sus parámetros).

(d) **[1.0 Ptos.]** ¿Cuál de los dos estimadores propuestos considera que es mejor? Justifique su respuesta.

Solución

(a) Tenemos que el valor esperado teórico de X es

$$\mathbf{E}(X) = \frac{\theta}{2} \quad \textbf{[0.5 Ptos.]}$$

Luego, el estimador de momentos de θ basado en las realizaciones de $x_1, \ldots, x_n$ de X es

$$\frac{\theta}{2} = \overline{x} \Rightarrow \hat{\theta} = 2\,\overline{x}, \quad \textbf{[0.5 Ptos.]}$$

con $\overline{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$.

(b) El estimador obtenido en (a) antes de ser evaluado en las realizaciones de la variable aleatoria X es una función de las variables aleatorias $X_1, \ldots, X_n$ cuya distribución es idéntica a la de X.

Se pide

$$\mathbf{E}(\hat{\theta}) = \mathbf{E}(2\,\overline{X}_n) = \frac{2}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \frac{2}{n}\sum_{i=1}^{n}\frac{\theta}{2} = \frac{2}{n}\cdot n\,\frac{\theta}{2} = \theta \quad \textbf{[1.0 Ptos.]}$$

Este resultado implica que $\hat{\theta}$ es un estimador insesgado para θ.

Por otra parte,

$$\begin{aligned}
\mathbf{Var}(\hat{\theta}) &= \mathbf{Var}(2\,\overline{X}_n) \\
&= \frac{4}{n^2}\mathbf{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{4}{n^2}\sum_{i=1}^{n}\mathbf{Var}\left(X_i\right), \quad \text{por independencia} \quad \textbf{[0.5 Ptos.]} \\
&= \frac{4}{n^2}\sum_{i=1}^{n}\frac{\theta^2}{12} \\
&= \frac{4}{n^2}\cdot n\,\frac{\theta^2}{12} \\
&= \frac{\theta^2}{3\,n} \quad \textbf{[0.5 Ptos.]}
\end{aligned}$$

(c) Consideremos el estimador máximo verosímil $\widetilde{\theta}$ de θ como función de las variables aleatorias $X_1, \ldots, X_n$ cuya distribución es idéntica a la de X.

$$\widetilde{\theta} = \max\{X_1, \ldots, X_n\}$$

La función de distribución de probabilidad acumulada de $\widetilde{\theta}$ está dada por

$$
\begin{aligned}
F_{\widetilde{\theta}}(u) &= P(\widetilde{\theta} \leq u) \\
&= P(\text{máx}\{X_1, \ldots, X_n\} \leq u) \\
&= P(X_1 \leq u, \ldots, X_n \leq u) \\
&= \prod_{i=1}^{n} P(X_i \leq u), \quad \text{por independencia} \quad \textbf{[0.5 Ptos.]} \\
&= F_X(u)^n, \quad \text{por idéntica distribución} \\
&= \left(\frac{u}{\theta}\right)^n \quad \textbf{[0.5 Ptos.]}
\end{aligned}
$$

Luego

$$
f_{\widetilde{\theta}}(u) = \frac{d}{du} F_{\widetilde{\theta}}(u) = n \left(\frac{u}{\theta}\right)^{n-1} \frac{1}{\theta}, \quad \textbf{[0.5 Ptos.]} \quad 0 \leq u \leq \theta \quad \textbf{[0.5 Ptos.]}
$$

(d) Los estimadores máximo verosímiles se caracterizan por ser asintóticamente insesgado de varianza mínima. Por esta razón, se prefieren ante cualquier otro estimador. **[1.0 Ptos.]**

+ 1 Punto Base

Problema 3

Hoy en día se discute la importancia de definir la matriz energética del futuro, de hecho hay estudios recientes que muestran que el gasto diario per cápita en tecnología (altamente demandante de energía eléctrica) alcanza a los \$4000 pesos y subiendo. Como es sabido, existe un grupo de ciudadanos (digamos "Verdes") preocupados del medio ambiente que se oponen a las represas (generadoras de electricidad). Sin embargo, personas de este grupo no son lo suficientemente coherentes entre el decir y el hacer. Más aún, es común declararse verde pero hacer uso (y abuso) de tecnología (digamos verdes de palabra). Pero hay otros verdes que dicen y hacen (digamos verdes comprometidos).

Un especialista afirma que

(a) "Hasta hoy en día, las personas catalogadas verdes son una minoría".

(b) "los catalogados como verdes tienen un gasto diario en tecnología inferior a la media nacional".

En cambio otro especialista afirma que

(c) "los verdes de palabra difieren en su gasto diario en tecnología con respecto a la media nacional".

Con el objetivo de contrastar los planteamientos anteriores usted lleva a cabo una encuesta a 120 personas, de las cuales 30 pueden ser catalogados como Verdes comprometidos, 20 como Verdes de palabras y los restantes como no Verdes. Al examinar sus gastos en tecnología, se tiene que los gastos diarios promedios son de \$3000, \$3500 y \$4500 pesos en cada grupo, respectivamente. A su vez, las desviaciones estándar de los gastos son de \$2300, \$1200 y \$1500 pesos respectivamente.

Con base a la información muestral obtenida y asumiendo normalidad en el gasto diario, lleve a cabo las pruebas de hipótesis necesarias que permitan contrastar las afirmaciones realizadas por los expertos en (a) **[2.0 Ptos.]**, (b) **[2.0 Ptos.]** y (c) **[2.0 Ptos.]**. Sea explícito, es decir, indique condiciones, test, prueba estadística, valor-p y conclusión para una probabilidad de error tipo I, α, igual a 5 %. Puede asumir normalidad en los gastos (y obviamente independencia).

Solución

(a) "Hasta hoy en día, las personas catalogadas verdes son una minoría". (Prueba de Hipótesis para una proporción) **[0.2 Ptos.]**

$$\mathrm{H}_0 : p = p_0 \quad \text{vs} \quad \mathrm{H}_a : p < p_0, \qquad \textbf{[0.5 Ptos.]}$$

con p la proporción de verdes (Comprometidos + De palabra) y $p_0 = 1/2$.

Para tamaños muestrales grandes, el estadístico de prueba

$$Z_n = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0\,(1 - p_0)}{n}}} \sim \text{Normal}(0, 1)$$

con

$$\hat{p} = \overline{X}_n = \frac{20 + 30}{120} = 0{,}4166667$$

Reemplazando

$$Z_n = \frac{0{,}41667 - 0{,}5}{\sqrt{0{,}5 \cdot 0{,}5/120}} = -1{,}825742 \quad \textbf{[0.5 Ptos.]}$$

La probabilidad de cometer un error tipo I esta dada por el valor-p,

$$\text{valor} - p = \Phi(-1{,}825742) = 1 - \Phi(1{,}825742) \approx 1 - \Phi(1{,}83) = 1 - 0{,}9664 = 0{,}0336 \quad \textbf{[0.5 Ptos.]}$$

Mientras que el valor crítico para una probabilidad "permitida", $\alpha = 5\,\%$, de cometer error tipo I es:

$$k_{0,05} = -k_{0,95} = -1{,}645$$

Alternativa 1

Por lo tanto, como el valor-$p < \alpha$, entonces existe evidencia suficiente que confirma la afirmación del especialista (Se rechaza H_0), es decir, los verdes son una minoría. **[0.3 Ptos.]**

Alternativa 2

Por lo tanto, como el estadístico de prueba Z_n es menor que al valor crítico, entonces existe evidencia suficiente que confirma la afirmación del especialista (Se rechaza H_0), es decir, los verdes son una minoría. **[0.3 Ptos.]**

(b) "los catalogados como verdes tienen un gasto diario en tecnología inferior a la media nacional". (Prueba de hipótesis para una media)

$$\text{Verdes} = \text{Verdes Comprometidos} + \text{Verdes de Paralabra}$$

Como los verder comprometidos son 30 y los de palabra son 20, el promedio de gasto diario en verdes se obtiene ponderando la información muestral de cada grupo, es decir,

$$\begin{aligned}\overline{X}_{\text{verde}} &= \frac{30}{50}\,\overline{X}_{\text{verde.c}} + \frac{20}{50}\,\overline{X}_{\text{verde.p}} \\ &= \frac{30}{50}\cdot 3000 + \frac{20}{50}\cdot 3500 \\ &= 3200 \quad \textbf{[0.2 Ptos.]}\end{aligned}$$

Mientras que la varianza muestral ponderada es

$$\begin{aligned}S^2_{\text{verde}} &= \frac{(30-1)\cdot S^2_{\text{verde.c}} + (20-1)\cdot S^2_{\text{verde.p}}}{50-1} \\ &= 3689184 \\ \Rightarrow S_{\text{verde}} &= 1920{,}725 \quad \textbf{[0.2 Ptos.]}\end{aligned}$$

Las hipótesis a testear son

$$\mathrm{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathrm{H}_a : \mu < \mu_0, \quad \textbf{[0.3 Ptos.]},$$

con $\mu_0 = 4000$.

Para tamaños muestrales grandes, el estadístico de prueba

$$T_n = \frac{\overline{X}_{\text{verde}} - \mu_0}{S_{\text{verde}}/\sqrt{n}} \sim \text{Normal}(0,1) \quad \textbf{[0.2 Ptos.]}$$

Reemplazando

$$T_n = \frac{3200 - 4000}{1920{,}725/\sqrt{50}} = -2{,}945166 \quad \textbf{[0.5 Ptos.]}$$

La probabilidad de cometer un error tipo I esta dada por el valor-p,

$$\text{valor} - p = \Phi(-2{,}945166) = 1 - \Phi(2{,}945166) \approx 1 - \Phi(2{,}95) = 1 - 0{,}9984 = 0{,}0016 \quad \textbf{[0.3 Ptos.]}$$

Alternativa 1

Por lo tanto, como el valor-$p < \alpha$, entonces existe evidencia suficiente que permite afirmar que los verdes tienen un gasto medio inferior a la media nacional (Se rechaza H_0). **[0.3 Ptos.]**

Alternativa 2

Por lo tanto, como el estadístico de prueba Z_n es menor que al valor crítico (-1.645), entonces existe evidencia suficiente que permite afirmar que los verdes tienen un gasto medio inferior a la media nacional (Se rechaza H_0). **[0.3 Ptos.]**

(c) "los verdes de palabra difieren en su gasto diario en tecnología con respecto a la media nacional". (Prueba de hipótesis para una media)

Las hipótesis a testear son

$$\mathrm{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathrm{H}_a : \mu \neq \mu_0, \quad \textbf{[0.3 Ptos.]},$$

con $\mu_0 = 4000$.

Para tamaños muestrales pequeños, el estadístico de prueba

$$T_n = \frac{\overline{X}_{\text{verde.p}} - \mu_0}{S_{\text{verde.p}}/\sqrt{n}} \sim t - \text{student}(n-1) \quad \textbf{[0.3 Ptos.]}$$

Reemplazando

$$T_n = \frac{3500 - 4000}{1200/\sqrt{20}} = -1{,}86339 \quad \textbf{[0.5 Ptos.]}$$

La probabilidad de cometer un error tipo I esta dada por el valor-p,

$$\text{valor} - p = 2\,P(T > |-1{,}86339|) = 2\,P(T > |-1{,}86339|) \Rightarrow 5\,\% < \text{valor} - p < 10\,\% \quad \textbf{[0.5 Ptos.]}$$

Mientras que el valor crítico para una probabilidad "permitida", $\alpha = 5\,\%$, de cometer error tipo I es:

$$t_{0{,}05/2}(19) = -t_{0{,}975}(19) = -2{,}093$$

Alternativa 1

Por lo tanto, como el valor-$p > \alpha$, entonces NO existe evidencia suficiente que que permite afirmar que los verdes de palabra difieran en su gasto medio con respecto a la media nacional (NO se rechaza H_0). **[0.4 Ptos.]**

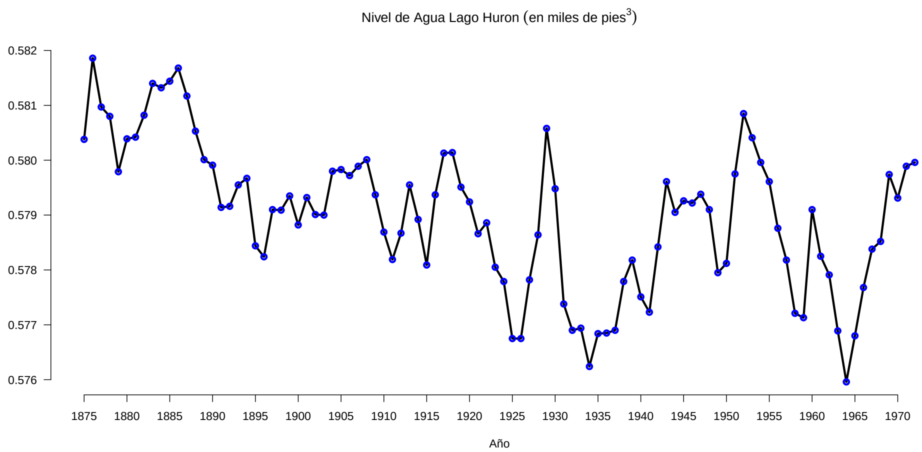
Alternativa 2

Por lo tanto, como el estadístico de prueba T_n es mayor que al valor crítico, entonces NO existe evidencia suficiente que que permite afirmar que los verdes de palabra difieran en su gasto medio con respecto a la media nacional (NO se rechaza H_0). **[0.4 Ptos.]**

+ 1 Punto Base

Problema 4

El lago Huron es el segundo lago más grande por área superficial y el quinto con agua dulce más grande del mundo. Un estudio realizado a mediados de los setenta considero los niveles anuales de agua entre los años 1875 y 1972 (Ver figura).



Un investigador piensa que en un año t, los niveles de agua X_t corresponde a una combinación lineal entre el nivel X_{t-1} y una variables aleatorias $\{Z_t\}$ independiente cuya distribución es Normal$(0,\,\sigma)$, es decir, el nivel X_t está dado por:

$$X_t = \nu + \phi\, X_{t-1} + Z_t,$$

con ν, ϕ **(con $|\phi| < 1$)** y σ parámetros.

Suponga que todas la medidas descriptiva asociadas a la variable aleatoria "nivel de agua" permanecen constantes en el tiempo, es decir, para todo tiempo t y s se tiene que

$$\mathbf{E}(X_t) = \mathbf{E}(X_s), \quad \mathbf{E}(X_t^2) = \mathbf{E}(X_s^2) \quad \text{y} \quad \mathbf{Cov}(X_t,\, X_{t+k}) = \mathbf{Cov}(X_s,\, X_{s+k}) \quad \text{para } k \in \mathbb{N}_0$$

(a) **[1.0 Ptos.]** Determine una expresión para el valor esperado, μ, del nivel de agua en el instante t, X_t, en términos de los parámetros ν y ϕ.

(b) **[2.0 Ptos.]** Muestre que $\mathbf{E}(X_t^2) = \dfrac{\nu^2}{(1-\phi)^2} + \dfrac{\sigma^2}{1-\phi^2}$.

(c) **[3.0 Ptos.]** Muestre que $\mathbf{Cov}(X_t,\, X_{t+k}) = \dfrac{\sigma^2\,\phi^k}{1-\phi^2}$, con $k \in \mathbb{N}_0$.

Solución

(a) Tenemos que

$$\begin{aligned}
\mathbf{E}(X_t) &= \mathbf{E}(\nu + \phi\, X_{t-1} + Z_t) \\
&= \nu + \phi\,\mathbf{E}(X_{t-1}) + \mathbf{E}(Z_t), \quad \text{por linealidad del operador esperanza} \quad \textbf{[0.3 Ptos.]} \\
\mu &= \nu + \phi\,\mu + 0, \quad \text{por enunciado} \quad \textbf{[0.3 Ptos.]}
\end{aligned}$$

Despejando, se tiene que

$$\mu = \frac{\nu}{1-\phi} \quad \textbf{[0.4 Ptos.]}$$

(b) Tenemos que

$$\begin{aligned}
\mathbf{E}(X_t^2) &= \mathbf{E}([\nu + \phi\, X_{t-1} + Z_t]^2) \quad \textbf{[0.2 Ptos.]} \\
&= E(\nu^2) + \phi^2\, \mathbf{E}(X_{t-1}^2) + \mathbf{E}(Z_t^2) + 2\,\nu\,\phi\, \mathbf{E}(X_{t-1}) + 2\,\nu\, \mathbf{E}(Z_t) + 2\,\phi\, \mathbf{E}(X_{t-1}\, Z_t) \quad \textbf{[0.2 Ptos.]} \\
&= E(\nu^2) + \phi^2\, \mathbf{E}(X_{t-1}^2) + \mathbf{E}(Z_t^2) + 2\,\nu\,\phi\, \mathbf{E}(X_{t-1}) + 2\,\nu\, \mathbf{E}(Z_t) + 2\,\phi\, \mathbf{E}(X_{t-1})\, \mathbf{E}(Z_t), \quad \text{por independencia} \quad \textbf{[0.3 Ptos.]} \\
&= \nu^2 + \phi^2\, \mathbf{E}(X_{t-1}^2) + \sigma^2 + 2\,\nu\,\phi\,\mu + 2\,\nu \cdot 0 + 2\,\phi\,\mu \cdot 0, \quad \textbf{[0.2 Ptos.]} \\
&= \nu^2 + \phi^2\, \mathbf{E}(X_t^2) + \sigma^2 + 2\,\nu\,\phi\,\mu, \quad \text{por enunciado} \quad \textbf{[0.2 Ptos.]}
\end{aligned}$$

Despejando, se tiene que

$$\begin{aligned}
E(X_t^2) &= \frac{\nu^2}{(1-\phi^2)} + \frac{\sigma^2}{(1-\phi^2)} + \frac{2\,\nu^2\,\phi}{(1-\phi)}\,\frac{1}{(1-\phi^2)} \quad \textbf{[0.2 Ptos.]} \\
&= \frac{\nu^2}{(1-\phi^2)\,(1-\phi)}\,[(1-\phi) + 2\,\phi] + \frac{\sigma^2}{(1-\phi^2)} \quad \textbf{[0.2 Ptos.]} \\
&= \frac{\nu^2}{(1-\phi)^2\,(1+\phi)}\,[1+\phi] + \frac{\sigma^2}{(1-\phi^2)} \quad \textbf{[0.3 Ptos.]} \\
&= \frac{\nu^2}{(1-\phi)^2} + \frac{\sigma^2}{(1-\phi^2)} \quad \textbf{[0.2 Ptos.]}
\end{aligned}$$

(c) Tenemos que

$$\begin{aligned}
X_t &= \nu + \phi\, X_{t-1} + Z_t \\
&= \nu + \phi\,(\nu + \phi\, X_{t-1} + Z_{t-1}) + Z_t \\
&= \nu + \phi\,\nu + \phi^2\, X_{t-1} + \phi\, Z_{t-1} + Z_t \\
&= \nu + \phi\,\nu + \phi^2\,(\nu + \phi\, X_{t-2} + Z_{t-2}) + \phi\, Z_{t-1} + Z_t \\
&= \nu + \phi\,\nu + \phi^2\,\nu + \phi^2\, X_{t-2} + \phi^2\, Z_{t-2} + \phi\, Z_{t-1} + Z_t \\
&\vdots \\
&= \nu\left(\sum_{j=0}^{\infty} \phi^j\right) + \left(\sum_{j=0}^{\infty} \phi^j\, Z_{t-j}\right) \quad \textbf{[0.2 Ptos.]} \\
&= \frac{\nu}{1-\phi} + \sum_{j=0}^{\infty} \phi^j\, Z_{t-j}, \quad \text{ya que } |\phi| < 1 \quad \textbf{[0.5 Ptos.]}
\end{aligned}$$

Se pide

$$\begin{aligned}
\mathbf{Cov}(X_t,\, X_{t+k}) &= \mathbf{Cov}\left(\frac{\nu}{1-\phi} + \sum_{j=0}^{\infty} \phi^j\, Z_{t-j},\, \frac{\nu}{1-\phi} + \sum_{i=0}^{\infty} \phi^i\, Z_{t+k-i}\right) \quad \textbf{[0.3 Ptos.]} \\
&= \mathbf{Cov}\left(\sum_{j=0}^{\infty} \phi^j\, Z_{t-j},\, \sum_{i=0}^{\infty} \phi^i\, Z_{t+k-i}\right), \quad \text{ya que } \frac{\nu}{1-\phi} \text{ es constante} \quad \textbf{[0.3 Ptos.]} \\
&= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \phi^{j+i}\, \mathbf{Cov}(Z_{t-j},\, Z_{t+k-i}), \quad \text{por linealidad del operador covarianza} \quad \textbf{[0.4 Ptos.]} \\
&= \sum_{j=0}^{\infty} \phi^{j+(k+j)}\, \sigma^2, \quad \text{ya que los } Z_t \text{ son independientes} \quad \textbf{[0.4 Ptos.]} \\
&= \sigma^2\, \phi^k \sum_{j=0}^{\infty} \phi^{2j} \quad \textbf{[0.2 Ptos.]}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{Cov}(X_t,\, X_{t+k}) &= \sigma^2\,\phi^k \sum_{j=0}^{\infty} (\phi^2)^j \quad \textcolor{red}{\textbf{[0.1 Ptos.]}} \\
&= \sigma^2\,\phi^k\,\frac{1}{1-\phi^2}, \quad \text{ya que } |\phi| < 1 \quad \textcolor{red}{\textbf{[0.3 Ptos.]}} \\
&= \frac{\sigma^2\,\phi^k}{1-\phi^2}, \quad \text{para } k \in \mathbb{N}_0 \quad \textcolor{red}{\textbf{[0.3 Ptos.]}}
\end{aligned}
$$

+ 1 Punto Base

Formulario

- Sea $X_1,\ldots,X_n$ una muestra aleatoria independiente e idénticamente distribuida con función de probabilidad p_X o de densidad f_X, determinada por un parámetro θ. Si $\hat{\theta}$ es el estimador máximo verosímil del parámetro θ, entonces:
  - $\mathbf{E}(\hat{\theta}) \to \theta$, cuando $n \to \infty$.
  - $\mathbf{Var}(\hat{\theta}) = \dfrac{1}{I_n(\theta)}$, con $I_n(\theta) = -\mathbf{E}\left[\dfrac{\partial^2}{\partial\,\theta^2}\,\ln L(\theta)\right]$.
  - $\hat{\theta} \,\dot{\sim}\, \text{Normal}\left(\theta,\, \sqrt{\dfrac{1}{I_n(\theta)}}\right)$, cuando $n \to \infty$.
  - El estimador máximo verosímil de $g(\theta)$ es $g(\hat{\theta})$, cuya varianza está dada por: $\mathbf{Var}[g(\hat{\theta})] = \dfrac{[g'(\theta)]^2}{I_n(\theta)}$.

Distribuciones

Distribución	Densidad de Probabilidad	Θ_X	Parámetros	Esperanza y Varianza
Binomial	$\binom{n}{x}\, p^x\,(1-p)^{n-x}$	$x = 0, \ldots, n$	$n,\ p$	$\mu_X = n\,p$ $\sigma_X^2 = n\,p\,(1-p)$
Geométrica	$p\,(1-p)^{x-1}$	$x = 1, 2, \ldots$	p	$\mu_X = 1/p$ $\sigma_X^2 = (1-p)/p^2$
Binomial-Negativa	$\binom{x-1}{r-1}\, p^r\,(1-p)^{x-r}$	$x = r, r+1, \ldots$	$r,\ p$	$\mu_X = r/p$ $\sigma_X^2 = r\,(1-p)/p^2$
Poisson	$\dfrac{(\nu\,t)^x\,e^{-\nu\,t}}{x!}$	$x = 0, 1, \ldots$	ν	$\mu_X = \nu\,t$ $\sigma_X^2 = \nu\,t$
Exponencial	$\nu\,e^{-\nu\,x}$	$x \geq 0$	ν	$\mu_X = 1/\nu$ $\sigma_X^2 = 1/\nu^2$
Gamma	$\dfrac{\nu^k}{\Gamma(k)}\,x^{k-1}\,e^{-\nu\,x}$	$x \geq 0$	$k,\ \nu$	$\mu_X = k/\nu$ $\sigma_X^2 = k/\nu^2$
Gamma Trasladada	$\dfrac{\nu^k}{\Gamma(k)}\,(x-\gamma)^{k-1}\,e^{-\nu\,(x-\gamma)}$	$x \geq \gamma$	$k,\ \nu,\ \gamma$	$\mu_X = k/\nu + \gamma$ $\sigma_X^2 = k/\nu^2$
Normal	$\dfrac{1}{\sqrt{2\,\pi}\,\sigma}\ \exp\left[-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right]$	$-\infty < x < \infty$	$\mu,\ \sigma$	$\mu_X = \mu$ $\sigma_X^2 = \sigma^2$
Log-Normal	$\dfrac{1}{\sqrt{2\,\pi}\,(\zeta\,x)}\ \exp\left[-\dfrac{1}{2}\left(\dfrac{\ln\,x-\lambda}{\zeta}\right)^2\right]$	$x \geq 0$	$\lambda,\ \zeta$	$\mu_X = \exp\left(\lambda + \frac{1}{2}\,\zeta^2\right)$ $\sigma_X^2 = \mu_X^2\left(e^{\zeta^2} - 1\right)$
Uniforme	$\dfrac{1}{(b-a)}$	$a \leq x \leq b$	$a,\ b$	$\mu_X = (a+b)/2$ $\sigma_X^2 = (b-a)^2/12$
Beta	$\dfrac{1}{B(q,\ r)}\ \dfrac{(x-a)^{q-1}\,(b-x)^{r-1}}{(b-a)^{q+r-1}}$	$a \leq x \leq b$	$q,\ r$	$\mu_X = a + \frac{q}{q+r}\,(b-a)$ $\sigma_X^2 = \frac{q\,r\,(b-a)^2}{(q+r)^2\,(q+r+1)}$
Hipergeométrica	$\dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$\text{máx}\{0, n+m-N\} \leq x \leq \text{mín}\{n, m\}$	$N,\ m,\ n$	$\mu_X = n\,\frac{m}{N}$ $\sigma_X^2 = \left(\frac{N-n}{N-1}\right)\,n\,\frac{m}{N}\,\left(1-\frac{m}{N}\right)$

- Propiedades función $\Gamma(\cdot)$:

$$(1)\quad \Gamma(k) = \int_0^\infty u^{k-1}\,e^{-u}\,du; \quad (2)\quad \Gamma(a+1) = a\,\Gamma(a);$$

$$(3)\quad \Gamma(n+1) = n!, \quad \text{si } n \in \mathbb{N}_0; \quad (4)\quad \Gamma(1/2) = \sqrt{\pi}$$

- Propiedades función $B(\cdot,\,\cdot)$:

$$(1)\quad B(q,\,r) = \int_0^1 x^{q-1}\,(1-x)^{r-1}\,dx; \quad (2)\quad B(q,\,r) = \frac{\Gamma(q)\,\Gamma(r)}{\Gamma(q+r)}$$

- Propiedad distribución Gamma:

$$\text{Si } T \sim \text{Gamma}(k,\,\nu) \Rightarrow F_T(t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu\,t)^x\,e^{-\nu\,t}}{x!}, \quad \text{si } k \in \mathbb{N}$$

Tablas de Percentiles p

Distribución Normal Estándar k_p

k_p	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	0,9925	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	0,9959	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
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Distribución t-student $t_p(\nu)$

ν	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$
1	3,078	6,314	12,706	31,821
2	1,886	2,920	4,303	6,965
3	1,638	2,353	3,182	4,541
4	1,533	2,132	2,776	3,747
5	1,476	2,015	2,571	3,365
6	1,440	1,943	2,447	3,143
7	1,415	1,895	2,365	2,998
8	1,397	1,860	2,306	2,896
9	1,383	1,833	2,262	2,821
10	1,372	1,812	2,228	2,764
11	1,363	1,796	2,201	2,718
12	1,356	1,782	2,179	2,681
13	1,350	1,771	2,160	2,650
14	1,345	1,761	2,145	2,624
15	1,341	1,753	2,131	2,602
16	1,337	1,746	2,120	2,583
17	1,333	1,740	2,110	2,567
18	1,330	1,734	2,101	2,552
19	1,328	1,729	2,093	2,539
20	1,325	1,725	2,086	2,528
21	1,323	1,721	2,080	2,518
22	1,321	1,717	2,074	2,508
23	1,319	1,714	2,069	2,500
24	1,318	1,711	2,064	2,492
25	1,316	1,708	2,060	2,485
26	1,315	1,706	2,056	2,479
27	1,314	1,703	2,052	2,473
28	1,313	1,701	2,048	2,467
29	1,311	1,699	2,045	2,462
30	1,310	1,697	2,042	2,457
∞	1,282	1,645	1,960	2,326

Distribución Chi-Cuadrado $c_p(\nu)$

ν	$c_{0,025}$	$c_{0,05}$	$c_{0,10}$	$c_{0,90}$	$c_{0,95}$	$c_{0,975}$	$c_{0,99}$	$c_{0,995}$
1	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	3,82	4,57	5,58	17,28	19,68	21,92	24,72	26,76
12	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
21	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40
22	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80
23	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18
24	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
25	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93