Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Segundo Semestre 2010

Curso	:	**Probabilidad y Estadística**
Sigla	:	**EYP1113**
Pauta	:	**I3**
Profesores	:	**Ricardo Aravena (Sec 02 y 03) y Ricardo Olea (Sec 01)**
Ayudantes	:	**Pablo Flores, Constanza Quezada, Victoria Veas.**

Problema 1

Considere una variable de interés (X) de una población, la cual está segmentada en k grupos. En el i-ésimo grupo, la variable X se rige por una distribución Normal cuyo valor esperado es $\mu\,\theta_i$, con θ_i una constante conocida para $i = 1, \ldots, k$. Si en cada uno de los k grupos se toma una muestra aleatoria de tamaño n_i, con $i = 1, \ldots, k$ y asumimos que la varianza en todos los grupos es constante e igual σ^2 (conocida).

(a) **[3.0 Ptos.]** Obtenga el estimador máximo verosímil de μ y muestre que su distribución es Normal. (Identifique los parámetros)

(b) **[3.0 Ptos.]** A partir de (a) proponga un intervalo de confianza de nivel $(1-\alpha)\cdot 100\,\%$ para μ.

Solución

(a) Sean $\{x_{i\,j}\}$ con $i = 1, \ldots, k$ y $j = 1, \ldots, n_i$ la muestra aleatoria tomada de la Población.

Tenemos que la función de verosimilitud está dada por:

$$
\begin{aligned}
L(\mu,\,\sigma) &= \prod_{i=1}^{k}\prod_{j=1}^{n_i} f_{X_i}(x_{i\,j}), \quad \text{por independencia}\\
&= \prod_{i=1}^{k}\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\,\pi\,\sigma^2}}\,\exp\left[-\frac{1}{2\,\sigma^2}(x_{i\,j}-\mu\,\theta_i)^2\right]\\
&= \left(\frac{1}{2\,\pi\,\sigma^2}\right)^{n/2}\exp\left[-\frac{1}{2\,\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{i\,j}-\mu\,\theta_i)^2\right], \quad \text{con } n=\sum_{i=1}^{k} n_i. \quad \textbf{[0.5 Ptos.]}
\end{aligned}
$$

Y la función log-verosimilitud queda como:

$$
\ell(\mu,\,\sigma) = -\frac{n}{2}\,\ln(2\,\pi) - \frac{n}{2}\,\ln(\sigma^2) - \frac{1}{2\,\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{i\,j}-\mu\,\theta_i)^2 \quad \textbf{[0.5 Ptos.]}
$$

Luego

$$
\frac{\partial}{\partial\,\mu}\,\ell(\mu,\,\sigma^2) = 0 \Rightarrow -\frac{1}{2\,\sigma^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i} 2\,(x_{i\,j}-\mu\,\theta_i)(-\theta_i) = 0
$$

Despejando, se tiene que

$$
\hat{\mu} = \frac{\displaystyle\sum_{i=1}^{k}\theta_i\left(\sum_{j=1}^{n_i} x_{i\,j}\right)}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} \quad \textbf{[0.5 Ptos.]}
$$

Su valor esperado y varianza son:

$$\mathbf{E}\left(\hat{\mu}\right) = \frac{\displaystyle\sum_{i=1}^{k} \theta_i \left(\sum_{j=1}^{n_i} \mathbf{E}(X_{i\,j})\right)}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} = \frac{\displaystyle\sum_{i=1}^{k} \theta_i \left(\sum_{j=1}^{n_i} \mu\,\theta_i\right)}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} = \frac{\displaystyle\mu\sum_{i=1}^{k} \theta_i \left(\sum_{j=1}^{n_i} \theta_i\right)}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2}$$

$$= \mu\,\frac{\displaystyle\sum_{i=1}^{k} \theta_i\,n_i\,\theta_i}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} = \mu\,\frac{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} = \mu \quad \textbf{[0.5 Ptos.]}$$

$$\mathbf{Var}\left(\hat{\mu}\right) = \frac{\displaystyle\sum_{i=1}^{k} \theta_i^2 \left(\sum_{j=1}^{n_i} \mathbf{Var}(X_{i\,j})\right)}{\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^2}, \quad \text{por independencia}$$

$$= \frac{\displaystyle\sum_{i=1}^{k} \theta_i^2 \left(\sum_{j=1}^{n_i} \sigma^2\right)}{\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^2} = \frac{\displaystyle\sum_{i=1}^{k} \theta_i^2\,n_i^2\,\sigma^2}{\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^2} = \sigma^2\,\frac{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2}{\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^2}$$

$$= \frac{\sigma^2}{\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]} \quad \textbf{[0.5 Ptos.]}$$

Dado que $\hat{\mu}$ es una combinación lineal de variables aleatorias con distribución Normal, la distribución Normal se mantiene **[0.5 Ptos.]**. Por lo tanto

$$\hat{\mu} \sim \text{Normal}\left(\mu,\,\sigma\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}\right)$$

con σ un valor conocido.

(b) De (a) tenemos que

$$\frac{\hat{\mu}-\mu}{\sigma\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}} \sim \text{Normal}(0,1) \quad \textbf{[0.5 Ptos.]}$$

Luego

$$P\left(k_{\alpha/2} \leq \frac{\hat{\mu}-\mu}{\sigma\displaystyle\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}} \leq k_{1-\alpha/2}\right) = 1-\alpha \quad \textbf{[0.5 Ptos.]}$$

$$\Rightarrow P\left(k_{\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \leq \hat{\mu}-\mu \leq k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}\right) = 1-\alpha \quad \textbf{[0.4 Ptos.]}$$

$$\Rightarrow P\left(-k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \leq \mu-\hat{\mu} \leq -k_{\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}\right) = 1-\alpha \quad \textbf{[0.4 Ptos.]}$$

$$\Rightarrow P\left(\hat{\mu}-k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \leq \mu \leq \hat{\mu}-k_{\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}\right) = 1-\alpha \quad \textbf{[0.4 Ptos.]}$$

$$\Rightarrow P\left(\hat{\mu}-k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \leq \mu \leq \hat{\mu}+k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2}\right) = 1-\alpha \quad \textbf{[0.4 Ptos.]}$$

Por lo tanto

$$<\mu>_{1-\alpha} \in \hat{\mu} \pm k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \qquad \textbf{[0.2 Ptos.]}$$

$$\in \frac{\displaystyle\sum_{i=1}^{k}\theta_i\left(\sum_{j=1}^{n_i} x_{i\,j}\right)}{\displaystyle\sum_{i=1}^{k} n_i\,\theta_i^2} \pm k_{1-\alpha/2}\,\sigma\,\left[\sum_{i=1}^{k} n_i\,\theta_i^2\right]^{-1/2} \qquad \textbf{[0.2 Ptos.]}$$

+ 1 Punto Base

Problema 2

Últimamente se ha discutido mediáticamente los gastos asociados a los diversos créditos que el mercado ofrece. Usted sostiene que la mayoría de las ofertas no incluyen en forma explícita el costo de los gastos asociados. Cómo una forma de demostrar su afirmación realiza una revisión de 50 "ofertas" (diferentes instituciones y/o diferentes tipos de créditos y/o diferentes montos), encontrando sólo en 20 de ellos la información referente a los gastos asociados. ¿Es evidencia suficiente para validar su afirmación? Use $\alpha = 10\,\%$.

Solución

Sea p la proporción de créditos que NO incluyen gastos asociados.

Del enunciado tenemos las siguientes hipótesis

$$\mathrm{H}_0 : p = p_0 \quad \text{vs} \quad \mathrm{H}_a : p > p_0, \quad \textbf{[2.0 Ptos.]}$$

con $p_0 = 0{,}5$.

El estadísticos de prueba bajo H_0

$$Z_n = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{\dfrac{30}{50} - 0{,}5}{\sqrt{\dfrac{0{,}5 \cdot (1-0{,}5)}{50}}} = 1{,}414 \quad \textbf{[2.0 Ptos.]}$$

distribuye aproximadamente Normal(0,1).

Alternativa 1

El valor-p está dado por:

$$\begin{aligned} \text{valor-p} &= P(Z > 1{,}414), \quad \text{con } Z \sim \text{Normal}(0,1) \\ &= 1 - P(Z \leq 1{,}414) \\ &\approx 1 - \Phi(1{,}41) \\ &= 1 - 0{,}9207 \\ &= 0{,}0793 \quad \textbf{[1.0 Ptos.]} \end{aligned}$$

Como el valor-p $= 7{,}93\,\% < 10\,\% = \alpha$, entonces se rechaza H_0, es decir, existe evidencia para fundamentar que la mayoría de las ofertas NO incluyen en forma explícita los cobros asociados. **[1.0 Ptos.]**

Alternativa 2

Para un nivel $\alpha = 10\,\%$ el valor crítico es $k_{1-\alpha} = 1{,}28$. **[1.0 Ptos.]**

Como $Z_n = 1{,}414 > 1{,}28 = k_{1-\alpha}$, entonces se rechaza H_0, es decir, existe evidencia para fundamentar que la mayoría de las ofertas NO incluyen en forma explícita los cobros asociados. **[1.0 Ptos.]**

+ 1 Punto Base

Problema 3

En un proceso de control de llenado de bebidas energéticas de 100 ml es de sobremanera relevante no producir un rebalse, es decir, el ideal es que las botellas contengan 100 ml. Por razón, cada cierto período se toman muestras de tamaño 20 y si el promedio de llenado es superior a 102 ml el proceso se detiene para efectuar una calibración. Suponga que el llenado de las botellas se comporta como una variable aleatoria con distribución Normal cuya desviación estándar σ es conocida e igual a 10 ml.

(a) **[2.0 Ptos.]** Determine la probabilidad de cometer un error tipo I.

(b) **[2.0 Ptos.]** Si el verdadero valor de μ es 103, determine la probabilidad de cometer error tipo II.

(c) **[2.0 Ptos.]** ¿Cuál debe ser el tamaño muestral para qué con la actual regla se obtenga un $\alpha = 5\,\%$ y un $\beta = 20\,\%$ cuando μ es igual a 103 ml?

Solución

(a) Las hipótesis a testear son

$$\mathrm{H}_0 : \mu = \mu_0 \quad \text{vs} \quad \mathrm{H}_a : \mu > \mu_0, \quad \textbf{[0.6 Ptos.]}$$

con $\mu_0 = 100$.

Bajo H_0 se tiene que

$$\frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} \sim \text{Normal}(0,1) \quad \textbf{[0.5 Ptos.]}$$

La probabilidad de error tipo I esta dada por:

$$\begin{aligned}
\alpha &= P(\text{Rechazar } \mathrm{H}_0 \,|\, \mathrm{H}_0 \text{ es verdadera}) \quad \textbf{[0.1 Ptos.]}\\
&= P(\overline{X}_n > 102 \,|\, \mu = \mu_0) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - P(\overline{X}_n \leq 102 \,|\, \mu = \mu_0) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - P\left(\frac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}} \leq \frac{102 - \mu_0}{\sigma/\sqrt{n}} \,\middle|\, \mu = \mu_0\right) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - \Phi\left(\frac{102 - 100}{10/\sqrt{20}}\right) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - \Phi\,(0{,}8944272) \quad \textbf{[0.1 Ptos.]}\\
&\approx 1 - \Phi\,(0{,}89) \quad \textbf{[0.1 Ptos.]}\\
&= 1 - 0{,}8133 \quad \textbf{[0.1 Ptos.]}\\
&= 0{,}1867 \quad \textbf{[0.1 Ptos.]}
\end{aligned}$$

(b) La probabilidad de error tipo II esta dada por

$$\begin{aligned}
\beta &= P(\text{No rechazar } \mathrm{H}_0 \,|\, \mathrm{H}_0 \text{ es falsa}) \quad \textbf{[0.2 Ptos.]}\\
&= P(\overline{X}_n \leq 102 \,|\, \mu = \mu_0 + \Delta) \quad \textbf{[0.4 Ptos.]}\\
&= P\left(\frac{\overline{X}_n - \mu_0 - \Delta}{\sigma/\sqrt{n}} \leq \frac{102 - \mu_0 - \Delta}{\sigma/\sqrt{n}} \,\middle|\, \mu = \mu_0 + \Delta\right) \quad \textbf{[0.2 Ptos.]}\\
&= \Phi\left(\frac{102 - 100 - 3}{10/\sqrt{20}}\right) \quad \textbf{[0.2 Ptos.]}\\
&= \Phi\,(-0{,}4472136) \quad \textbf{[0.2 Ptos.]}\\
&\approx \Phi\,(-0{,}45) \quad \textbf{[0.2 Ptos.]}\\
&= 1 - \Phi(0{,}45) \quad \textbf{[0.2 Ptos.]}\\
&= 1 - 0{,}6736 \quad \textbf{[0.2 Ptos.]}\\
&= 0{,}3264 \quad \textbf{[0.2 Ptos.]}
\end{aligned}$$

(c) Para un error tipo I, α, la regla de rechazo bajo H_0 es

$$\overline{X}_n > \mu_0 + k_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \quad \textbf{[0.3 Ptos.]}$$

El error tipo II, controlado por el error tipo I esta dado por

$$\begin{aligned}
\beta &= P\left(\overline{X}_n \leq \mu_0 + k_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \,\middle|\, \mu = \mu_0 + \Delta\right) \quad \textbf{[0.2 Ptos.]} \\
&= P\left(\frac{\overline{X}_n - \mu_0 - \Delta}{\sigma/\sqrt{n}} \leq k_{1-\alpha} - \frac{\Delta\sqrt{n}}{\sigma} \,\middle|\, \mu = \mu_0 + \Delta\right) \quad \textbf{[0.2 Ptos.]} \\
&= \Phi\left(k_{1-\alpha} - \frac{\Delta\sqrt{n}}{\sigma}\right) \Rightarrow k_\beta = k_{1-\alpha} - \frac{\Delta\sqrt{n}}{\sigma} \quad \textbf{[0.2 Ptos.]}
\end{aligned}$$

Despejando se tiene que

$$n = \left(\frac{\sigma\,[k_{1-\alpha} - k_\beta]}{\Delta}\right)^2 = \left(\frac{\sigma\,[k_{1-\alpha} + k_{1-\beta}]}{\Delta}\right)^2 \quad \textbf{[0.3 Ptos.]}$$

donde $k_{1-\alpha} = k_{1-0{,}05} \approx 1{,}64$, $k_{1-\beta} = k_{1-0{,}20} \approx 0{,}84$ y $\Delta = 100 - 103 = 3$. **[0.3 Ptos.]**

Reemplazando

$$n = \left(\frac{10 \cdot (1{,}64 + 0{,}84)}{3}\right)^2 = 68{,}33778 \quad \textbf{[0.3 Ptos.]}$$

Por lo tanto el tamaño muestral debe ser al menos 69 unidades. **[0.2 Ptos.]**

+ 1 Punto Base

Problema 4

Durante un día hábil se han registrados los tiempos entre llegadas de clientes a una tienda. El encargado de llevar el registro consideró que era mejor resumir la información y entregó los tiempos transcurridos cada dos clientes. Sean $t_1, \ldots, t_n$ los primeros n tiempos entregados.

(a) **[3.0 Ptos.]** Proponga una distribución adecuada para los tiempos entregados y estime el(los) parámetro(s) por el método de máxima verosimilitud.

(b) **[3.0 Ptos.]** Si la media muestral de los primeros 30 tiempos entregados fue de 3 minutos con una desviación estándar de 1,5 minutos y el resumen de los datos es el siguiente:

Intervalo de Tiempo (Minutos)	[0, 3)	[3, 6)	[6, 9)	[9, $+\infty$)
Frecuencia	19	6	3	2

Evalúe la calidad de ajuste que tienen los tiempo con respecto a la distribución propuesta en (a). Use $\alpha = 5\,\%$.

Solución

(a) Tenemos que un proceso adecuado para modelar las llegadas de clientes en un intervalo de tiempo $[0, t]$ a una tienda es el proceso de Poisson($\nu\, t$). Los tiempos entre llegadas de clientes para este proceso distribuyen independientemente Exponencial(ν) y los tiempos transcurridos cada k clientes tiene una distribución Gamma(k, ν). **[0.5 Ptos.]**

Para el caso planteado, los tiempos entregados $t_1, \ldots, t_n$ son realizaciones de variables aleatorias independientes cuya distribución es Gamma($k = 2$, ν). **[1.0 Ptos.]**

La función de verosimilitud esta dada por

$$\begin{aligned} L(\nu) &= \prod_{i=1}^{n} f_{T_i}(t_i), \quad \text{por independencia} \\ &= \prod_{i=1}^{n} \nu^2\, t_i\, \exp(-\nu\, t_i) \\ &= \nu^{2\,n} \left[\prod_{i=1}^{n} t_i\right] \exp\left(-\nu \sum_{i=1}^{n} t_i\right) \quad \textbf{[0.5 Ptos.]} \end{aligned}$$

Mientras que la log-verosimilitud queda como

$$\ell(\nu) = 2\,n\, \ln(\nu) + \sum_{i=1}^{n} \ln(t_i) - \nu \sum_{i=1}^{n} t_i \quad \textbf{[0.5 Ptos.]}$$

Luego

$$\frac{\partial}{\partial\,\nu}\,\ell(\nu) = 0 \Rightarrow \frac{2\,n}{\nu} - \sum_{i=1}^{n} t_i = 0$$

Despejando, se tiene que

$$\hat{\nu} = \frac{2}{\bar{t}}, \quad \text{con } \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i \quad \textbf{[0.5 Ptos.]}$$

(b) Como el promedio de los tiempos entregados fue 3 minutos, entonces la distribución teórica estimada es una Gamma($k = 2$, $\nu = 2/3$). **[0.4 Ptos.]**

Luego, las probabilidades que una variable aleatoria T cuya distribución de probabilidad es Gamma(2, 2/3) se encuentre en cada uno de los cuatro intervalos que presentan son:

$$p_1 = P(0 \leq T < 3) = F_T(3) \quad \textbf{[0.1 Ptos.]}$$
$$p_2 = P(3 \leq T < 6) = F_T(6) - F_T(3) \quad \textbf{[0.1 Ptos.]}$$
$$p_3 = P(6 \leq T < 9) = F_T(9) - F_T(6) \quad \textbf{[0.1 Ptos.]}$$
$$p_4 = P(9 \leq T) = 1 - F_T(9) \quad \textbf{[0.1 Ptos.]}$$

Del formulario tenemos que

$$F_T(t) = 1 - \sum_{x=0}^{2-1} \frac{(\nu\, t)^x\, e^{-\nu\, t}}{x!} = 1 - e^{-2\,t/3} - (2\,t/3)\, e^{-2\,t/3} \quad \textbf{[0.2 Ptos.]}$$

Esto implica que

$$F_T(3) = 1 - e^{-2} - 2\, e^{-2} = 0{,}5939942$$
$$F_T(6) = 1 - e^{-4} - 4\, e^{-4} = 0{,}9084218 \quad \textbf{[0.3 Ptos.]}$$
$$F_T(9) = 1 - e^{-6} - 6\, e^{-6} = 0{,}9826487$$

Reemplazando

$$p_1 = 0{,}5939942; \quad p_2 = 0{,}3144276; \quad p_3 = 0{,}0742269; \quad p_4 = 0{,}0173513 \quad \textbf{[0.3 Ptos.]}$$

Por lo tanto, los valores esperados y observados quedan como sigue:

Intervalo de Tiempo (Minutos)	[0, 3)	[3, 6)	[6, 9)	[9, $+\infty$)
Observado	19	6	3	2
Esperado	17,819826	9,432828	2,226807	0,520539

[0.4 Ptos.]

El estadístico de prueba

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 5{,}800798 \quad \textbf{[0.3 Ptos.]}$$

tiene una distribución χ^2_f, con $f = (k - 1) - 1 = 2$. **[0.3 Ptos.]**

Alternativa 1

La probabilidad de Error Tipo I asociada al estadístico de prueba se encuentra en el intervalo

$$5\,\% < \text{valor-p} < 10\,\% \quad \textbf{[0.2 Ptos.]}$$

Por lo tanto, como el valor-p es mayor que el nivel α igual al 5 %, la hipótesis que los tiempos entregados distribuyen Gamma(2, 2/3) no es rechazada. **[0.2 Ptos.]**

Alternativa 2

El valor crítico para un nivel $\alpha = 5\,\%$ es $c_{1-0{,}05}(2) = 5{,}99$. **[0.2 Ptos.]**

Por lo tanto, como $X^2 = 5{,}800798 < 5{,}99 = c_{1-0{,}05}(2)$, la hipótesis que los tiempos entregados distribuyen Gamma(2, 2/3) no es rechazada. **[0.2 Ptos.]**

+ 1 Punto Base

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	$\dfrac{1}{\sqrt{2\,\pi}\,(\zeta\,x)}\,\exp\left[-\dfrac{1}{2}\left(\dfrac{\ln\,x-\lambda}{\zeta}\right)^2\right]$	$x \geq 0$	$\lambda,\ \zeta$	$\mu_X = \exp\left(\lambda + \dfrac{1}{2}\,\zeta^2\right)$ $\sigma_X^2 = \mu_X^2\left(e^{\zeta^2} - 1\right)$
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	$\dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$	$\max\{0, n+m-N\} \leq x$ $x \leq \min\{n, m\}$	$N,\ m,\ n$	$\mu_X = n\,\frac{m}{N}$ $\sigma_X^2 = \left(\frac{N-n}{N-1}\right)\,n\,\frac{m}{N}\,\left(1-\frac{m}{N}\right)$

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

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

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

- 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

Problema 4 (ALTERNATIVA)

Durante un día hábil se han registrados los tiempos entre llegadas de clientes a una tienda. El encargado de llevar el registro consideró que era mejor resumir la información y entregó los tiempos transcurridos cada dos clientes. Sean $t_1, \ldots, t_n$ los primeros n tiempos entregados.

(a) **[3.0 Ptos.]** Proponga una distribución adecuada para los tiempos entregados y estime el(los) parámetro(s) por el método de máxima verosimilitud.

(b) **[3.0 Ptos.]** Si la media muestral de los primeros 30 tiempos entregados fue de 3 minutos con una desviación estándar de 1,5 minutos y el resumen de los datos es el siguiente:

Intervalo de Tiempo (Minutos)	[0, 3)	[3, 6)	[6, 9)	[9, $+\infty$)
Frecuencia	19	6	3	2

Evalúe la calidad de ajuste que tienen los tiempo con respecto a la distribución propuesta en (a). Use $\alpha = 5\,\%$.

Solución

(a) Tenemos que un proceso adecuado para modelar las llegadas de clientes en un intervalo de tiempo $[0,\, t]$ a una tienda es el proceso de Poisson($\nu\, t$). Los tiempos entre llegadas de clientes para este proceso distribuyen independientemente Exponencial(ν) y los tiempos transcurridos cada k clientes también tiene una distribución Exponencial(ν). **[0.0 Ptos.]**

Para el caso planteado, los tiempos entregados $t_1, \ldots, t_n$ son realizaciones de variables aleatorias independientes cuya distribución es Exponencial(ν) = Gamma($k = 1,\, \nu$). **[0.0 Ptos.]**

La función de verosimilitud esta dada por

$$
\begin{aligned}
L(\nu) &= \prod_{i=1}^{n} f_{T_i}(t_i), \quad \text{por independencia} \\
&= \prod_{i=1}^{n} \nu \, \exp\left(-\nu\, t_i\right) \\
&= \nu^n \, \exp\left(-\nu \sum_{i=1}^{n} t_i\right) \qquad \textbf{[0.3 Ptos.]}
\end{aligned}
$$

Mientras que la log-verosimilitud queda como

$$\ell(\nu) = n\, \ln(\nu) - \nu \sum_{i=1}^{n} t_i \quad \textbf{[0.3 Ptos.]}$$

Luego

$$\frac{\partial}{\partial\, \nu}\, \ell(\nu) = 0 \Rightarrow \frac{n}{\nu} - \sum_{i=1}^{n} t_i = 0$$

Despejando, se tiene que

$$\hat{\nu} = \frac{1}{\bar{t}}, \quad \text{con } \bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i \quad \textbf{[0.4 Ptos.]}$$

(b) Como el promedio de los tiempos entregados fue 3 minutos, entonces la distribución teórica estimada es una Gamma($k = 1$, $\nu = 1/3$). **[0.4 Ptos.]**

Luego, las probabilidades que una variable aleatoria T cuya distribución de probabilidad es Gamma(1, 1/3) se encuentre en cada uno de los cuatro intervalos que presentan son:

$$p_1 = P(0 \leq T < 3) = F_T(3) \quad \textbf{[0.1 Ptos.]}$$
$$p_2 = P(3 \leq T < 6) = F_T(6) - F_T(3) \quad \textbf{[0.1 Ptos.]}$$
$$p_3 = P(6 \leq T < 9) = F_T(9) - F_T(6) \quad \textbf{[0.1 Ptos.]}$$
$$p_4 = P(9 \leq T) = 1 - F_T(9) \quad \textbf{[0.1 Ptos.]}$$

Del formulario tenemos que

$$F_T(t) = 1 - \sum_{x=0}^{1-1} \frac{(\nu\, t)^x\, e^{-\nu\, t}}{x!} = 1 - e^{-t/3} \quad \textbf{[0.2 Ptos.]}$$

Esto implica que

$$F_T(3) = 1 - e^{-2} - 2\, e^{-2} = 0{,}6321206$$
$$F_T(6) = 1 - e^{-4} - 4\, e^{-4} = 0{,}8646647 \quad \textbf{[0.3 Ptos.]}$$
$$F_T(9) = 1 - e^{-6} - 6\, e^{-6} = 0{,}950213$$

Reemplazando

$$p_1 = 0{,}63212056; \quad p_2 = 0{,}23254416; \quad p_3 = 0{,}08554821; \quad p_4 = 0{,}04978707 \quad \textbf{[0.3 Ptos.]}$$

Por lo tanto, los valores esperados y observados quedan como sigue:

Intervalo de Tiempo (Minutos)	[0, 3)	[3, 6)	[6, 9)	[9, $+\infty$)
Observado	19	6	3	2
Esperado	18,963617	6,976325	2,566446	1,493612

[0.4 Ptos.]

El estadístico de prueba

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = 0{,}3816293 \quad \textbf{[0.3 Ptos.]}$$

tiene una distribución χ^2_f, con $f = (k - 1) - 1 = 2$. **[0.3 Ptos.]**

Alternativa 1

La probabilidad de Error Tipo I asociada al estadístico de prueba se encuentra en el intervalo

$$10\,\% < \text{valor-p} < 90\,\% \quad \textbf{[0.2 Ptos.]}$$

Por lo tanto, como el valor-p es mayor que el nivel α igual al 5 %, la hipótesis que los tiempos entregados distribuyen Gamma(1, 1/3) no es rechazada. **[0.2 Ptos.]**

Alternativa 2

El valor crítico para un nivel $\alpha = 5\,\%$ es $c_{1-0{,}05}(2) = 5{,}99$. **[0.2 Ptos.]**

Por lo tanto, como $X^2 = 0{,}3816293 < 5{,}99 = c_{1-0{,}05}(2)$, la hipótesis que los tiempos entregados distribuyen Gamma(1, 1/3) no es rechazada. **[0.2 Ptos.]**

+ 1 Punto Base