Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Estadística

Primer Semestre 2025

Curso : Probabilidad y Estadística
Sigla : EYP1113
Profesores : Ricardo Aravena C., Maritza Márquez N., Bladimir Morales T., Ricardo Olea O., Jose Ordoñez C. y Hernán Robledo A.

INTERROGACIÓN 2

Pregunta 1

En marzo de 2025, el secretario de Defensa de EE.UU., Pete Hegseth, compartió por error información sensible sobre Yemen en un chat de Signal que incluía a un periodista. Este incidente, conocido como "Signalgate", generó preocupación sobre la posibilidad de que mensajes enviados en grupos cerrados contengan información comprometedora.

Para modelar este tipo de situaciones, se supone que varios funcionarios envían mensajes en el grupo, y que cada mensaje tiene una probabilidad $\mathcal{P}$ de incluir información sensible. Como distintos miembros pueden tener diferentes prácticas de seguridad, se modela la incertidumbre sobre $\mathcal{P}$ mediante una distribución Beta$(1, 2)$, lo que refleja que es más probable que un mensaje no contenga información sensible, aunque no se descarta el riesgo.

Suponga que se enviaron 7 mensajes recientes, cada uno por un funcionario distinto. ¿Cuál es la probabilidad de que exactamente 2 de estos mensajes hayan incluido información confidencial?

Solución

La probabilidad de que un mensaje incluya información confidencial está dada por una variable aleatoria $\mathcal{P} \sim$ Beta$(1, 2)$. Condicional a un valor fijo de $\mathcal{P} = p$, el número de mensajes con información confidencial entre los $n = 7$ mensajes enviados sigue una distribución binomial $X \mid \mathcal{P} = p \sim$ Binomial$(n = 7, p)$.

Se requiere entonces la probabilidad marginal de que exactamente 2 mensajes hayan contenido información confidencial:

$$P(X = 2) = \int_0^1 P(X = 2 \mid \mathcal{P} = p) f_{\mathcal{P}}(p)\, dp$$

Con $f_{\mathcal{P}}(p) = p^{1-1}(1-p)^{2-1}/B(1,2) = 2(1-p)$ para $p \in [0, 1]$, y la distribución binomial condicional:

$$P(X = 2 \mid \mathcal{P} = p) = \binom{7}{2} p^2 (1-p)^5$$

Entonces,

$$\begin{aligned} P(X = 2) &= \int_0^1 \binom{7}{2} p^2 (1-p)^5 \cdot 2(1-p)\, dp \\ &= \binom{7}{2} \cdot 2 \int_0^1 p^2 (1-p)^6\, dp \\ &= 21 \cdot 2 \cdot \int_0^1 p^2 (1-p)^6\, dp \\ &= 42 \cdot B(3, 7) \end{aligned}$$

Donde $B(a,b)$ es la función beta, con:

$$B(3,7) = \frac{\Gamma(3)\Gamma(7)}{\Gamma(10)} = \frac{2! \cdot 6!}{9!} = \frac{2 \cdot 720}{362880} = \frac{1440}{362880} = \frac{1}{252}$$

Por lo tanto:

$$P(X=2) = 42 \cdot \frac{1}{252} = \frac{1}{6} \approx 0.1667$$

Asignación de Puntaje:

Logro 1 [2 ptos]: por definir variables aleatorias del enunciado.

- [0.5 pto] por indicar que $\mathcal{P} \sim$ Beta$(1,2)$.
- [1 pto] por indicar que $X|\mathcal{P}=p \sim$ Binomial$(7,p)$
- [0.5 pto] por notar que se solicita $P(X=2)$ o $P(X=x)$.

Logro 2 [2 ptos]: por construir correctamente la integral de probabilidad marginal.

- [2 pto] por plantear $P \int_0^1 P(X=x \mid \mathcal{P}=p) f_{\mathcal{P}}(p)\, dp$.

Logro 3 [2 ptos]: por resolver la integral y obtener el resultado final.

- [1 pto] por desarrollar la integral y obtener que es igual a $B(3,7)$.
- [1 pto] por obtener $B(3,7) = 1/252$ y concluir $P(X=2) = 1/6$.

Pregunta 2

En mayo de 2025, el Centro Meteorológico Marítimo de Valparaíso emitió un aviso de marejadas que afectó extensamente el litoral chileno, desde el Golfo de Penas hasta Arica, incluyendo el Archipiélago Juan Fernández. Estas marejadas, generadas por un sistema frontal en el sur del país, alcanzaron su mayor intensidad durante las horas de pleamar (marea alta).

Para fines de planificación en ingeniería civil costera, se busca modelar la frecuencia e intensidad de las "marejadas severas", entendidas como aquellas marejadas cuyo promedio de alturas de olas supera los 4 metros.

(a) **[2.0 Ptos.]** Suponga que la ocurrrencia de marejadas sigue un proceso de Poisson con una tasa promedio de 3 marejadas por día. ¿Cuál es la probabilidad de que en 2 días ocurran exactamente 5 marejadas?

(b) **[2.0 Ptos.]** ¿Cuál es la probabilidad de que el tiempo entre dos marejadas sea mayor a 12 horas?

(c) **[2.0 Ptos.]** Considere que, en cada marejada, el promedio de las alturas de las olas (en metros) sigue una distribución Gamma con coeficiente de variación igual a $1/\sqrt{2}$ y un valor esperado de 2.5 metros. ¿Cuántos días se espera que pasen entre tres marejadas severas?

Solución

(a) Si las marejadas siguen un proceso de Poisson con tasa $\lambda = 3$ marejadas por día, entonces en un intervalo de 2 días, la tasa esperada es $\lambda_t = 2 \cdot 3 = 6$. Así, la variable aleatoria X que cuenta la cantidad de marejadas en 2 días se distribuye como $X_2 \sim$ Poisson(6). Se pide $P(X_2 = 5)$:

$$P(X_2 = 5) = \frac{6^5}{5!}e^{-6} = \frac{7776}{120}e^{-6} = 64.8 \cdot e^{-6} \approx 64.8 \cdot 0.00247875 \approx 0.1606$$

(b) En un proceso de Poisson con tasa $\lambda = 3$ por día, el tiempo entre eventos sigue una distribución Exponencial con parámetro $\lambda = 3$. Convertimos 12 horas a días: 12 hrs $= 0.5$ días. Entonces, la probabilidad de que el tiempo entre marejadas sea mayor a 0.5 días es:

$$P(T > 0.5) = e^{-\lambda t} = e^{-3 \cdot 0.5} = e^{-1.5} \approx 0.2231$$

(c) Como la ocurrencia de marejadas sigue un proceso Poisson, y las marejadas severas son una proporción p de estas, donde p corresponde a la probabilidad de que el promedio de alturas de olas supere los 4 metros, entonces la ocurrencia de marejadas severas también sigue un proceso Poisson de tasa $\lambda_t \cdot p$.

Se pide el tiempo esperado en días entre tres marejadas severas, el cual corresponde a $\mathbb{E}(T_2)$, donde $T_2 \sim$ Gamma$(2, \lambda_t \cdot p)$ por ser parte del proceso Poisson. Como $\lambda_t = 3$, procedemos a calcular p.

Sea Y el promedio de alturas de olas en una marejada, con $Y \sim$ Gamma(k, ν). Los parámetros k y ν se obtienen a partir de los datos del enunciado como sigue:

$$\mathbb{E}(Y) = \frac{k}{\nu} = 2.5, \quad \frac{1}{2} = \delta = \frac{\sigma}{\mu} = \frac{\sqrt{k/\nu^2}}{k/\nu} = \frac{1}{\sqrt{k}} \quad \Longrightarrow \quad k = 2$$

Así,

$$\mathbb{E}(Y) = \frac{2}{\nu} = 2.5 \quad \Longrightarrow \quad \nu = \frac{2}{2.5} = 0.8.$$

Necesitamos calcular $p = P(Y > 4)$. Del formulario se tiene que si $Y \sim$ Gamma(k, ν) con $k \in \mathbb{N}$, entonces:

$$F_Y(t) = P(Y < t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu t)^x e^{-\nu t}}{x!}$$

Reemplazando $k = 2$ y $\nu = 0.8$:

$$P(Y > 4) = 1 - F_Y(4) = \sum_{x=0}^{1} \frac{(0.8 \cdot 4)^x e^{-0.8 \cdot 4}}{x!} = \frac{(3.2)^0 e^{-3.2}}{0!} + \frac{(3.2)^1 e^{-3.2}}{1!} = e^{-3.2}(1 + 3.2)$$

Así,

$$e^{-3.2} \approx 0.04076 \quad \Longrightarrow \quad p = P(Y > 4) \approx 0.04076 \cdot 4.2 \approx 0.1712$$

Finalmente, el número esperado entre tres marejadas:

$$\mathbb{E}(T_2) = \frac{2}{\lambda_t \cdot p} = \frac{2}{3 \cdot 0.1712} = 3.894081 \text{ días.}$$

***Observación**: En las aclaraciones se puso que se podía interpretar como el tiempo esperado hasta la tercera marejada severa, lo que cambia la respuesta a $\mathbb{E}(T_3) = 3/(3 \cdot 0.1712) = 5.841$. Este desarrollo será considerado como correcto.*

Asignación de Puntaje:

a)

Logro 1 [1 ptos]: por transformar tasa del proceso a dos días como $2 \cdot \lambda_t = 6$.
Logro 2 [1 ptos]: por calcular correctamente $P(X = 5) = 0.1606$.

b)

Logro 3 [1 ptos]: por reconocer que el tiempo entre eventos sigue una distribución Exponencial con $\lambda = 3$.
Logro 4 [1 ptos]: por calcular $P(T > 0.5) = e^{-1.5} \approx 0.2231$.

- [0.5 ptos] por reconocer que se pide $P(T > 0.5)$.
- [0.5 ptos] por desarrollar y obtener $P(T > 0.5) = 0.2231$.

c)

Logro 5 [1 ptos]: por reconocer lo solicitado: $\mathbb{E}(T_2)$, con $T_2 \sim$ Gamma$(2, \lambda_t \cdot p)$.

- [0.3 ptos] por reconocer que las marejadas severas también siguen un proceso Poisson de tasa $\lambda_t \cdot p$.
- [0.2 ptos] por reconocer que $p = P(Y > 4)$.
- [0.5 ptos] por indicar que por ser proceso Poisson, $T_3 \sim$ Gamma$(3, \lambda_t \cdot p)$.

Logro 6 [1 ptos]: por obtener lo necesario para obtener $\mathbb{E}(T_2)$:

- [0.5 ptos] por obtener parámetros $k = 2$ y $\nu = 0.8$.
- [0.3 ptos] por calcular $P(Y > 4) = 0.1712$.
- [0.2 ptos] por obtener $\mathbb{E}(T_2) = 3.8940$(o 5.841 en el caso que considero el tiempo hasta la tercera marejada severa).

Pregunta 3

En São Paulo, la expansión de la droga sintética K9, conocida como la "droga zombi", ha afectado gravemente a personas en situación de calle. Un grupo de investigadores en toxicología realizó un estudio para caracterizar clínicamente sus efectos, recopilando datos sobre tres aspectos clave: el tiempo hasta la aparición de síntomas severos tras la inhalación, el tiempo hasta recibir atención médica, y la duración de conductas violentas provocadas por la sustancia. El análisis de los datos arrojó los siguientes resultados:

- El tiempo hasta la aparición de síntomas severos sigue una distribución Log-Normal, con mediana de 12 minutos y coeficiente de variación del 45 %.
- El tiempo entre la aparición de los síntomas y la atención médica en un centro de salud sigue una distribución Weibull, con mediana de 45 minutos y percentil 60 % igual a 50 minutos.
- La duración de las conductas violentas sigue una distribución Logística, con media de duración de 15 minutos y desviación estándar de 7 minutos.

Responda las siguientes preguntas de investigación:

(a) **[2.0 Ptos.]** ¿Cuál es la probabilidad de que una persona demore más de 24 minutos en presentar síntomas severos?

(b) **[2.0 Ptos.]** ¿Cuál es la probabilidad de que el tiempo hasta recibir atención médica sea mayor a una hora?

(c) **[2.0 Ptos.]** ¿Cuál es la probabilidad de que las conductas violentas sean menores a 5 minutos?

Solución

(a) Sea X el tiempo hasta la aparición de síntomas, con $X \sim$ LogNormal(λ, ξ). Del enunciado:

- Mediana: $e^{\lambda} = 12 \Rightarrow \lambda = \log(12)$
- Coeficiente de variación: $\delta = \sqrt{e^{\xi^2} - 1} = 0.45 \Rightarrow \xi^2 = \log(1 + 0.45^2) = \log(1.2025) \approx 0.184$. Entonces, $\xi = \sqrt{0.184} \approx 0.429$

Buscamos $P(X > 24)$:

$$\begin{aligned} P(X > 24) &= P\left(\log(X) > \log(24)\right) = P\left(Z > \frac{\log(24) - \log(12)}{0.429}\right) \\ &= P\left(Z > \frac{\log(2)}{0.429}\right) = P(Z > 1.616) \\ &= 1 - \Phi(1.616) \approx 1 - 0.947 = 0.053 \end{aligned}$$

(b) Sea $Y \sim$ Weibull(η, β). Del formulario se tiene que:

$$F_Y(t) = 1 - \exp\left(-\left(\frac{t}{\eta}\right)^{\beta}\right),$$

por lo que al despejar se obtiene que un percentil t_p es:

$$t_p = \eta\left(-\log(1-p)\right)^{1/\beta}.$$

Así, la mediana es $t_{0.5} = \eta\left(-\log(0.5)\right)^{1/\beta} = 45$ y el percentil 60 % es $t_{0.6} = \eta\left(-\log(0.4)\right)^{1/\beta} = 50$.. Para despejar el parámetro β, dividimos las dos expresiones:

$$\frac{t_{0.6}}{t_{0.5}} = \frac{50}{45} = \left(\frac{-\log(0.4)}{-\log(0.5)}\right)^{1/\beta} \Rightarrow \frac{10}{9} = (1.322)^{1/\beta} \Rightarrow \beta = \frac{\log(1.322)}{\log(10/9)} \approx \frac{0.279}{0.105} \approx 2.649.$$

Reemplazamos en la fórmula de la mediana para obtener η:

$$45 = \eta\,(\log 2)^{1/2.649} \Rightarrow \eta = \frac{45}{(\log 2)^{1/2.649}} \approx \frac{45}{0.933} \approx 51.67.$$

Finalmente, calculamos lo solicitado:

$$P(Y > 60) = \exp\left(-\left(\frac{60}{\eta}\right)^{\beta}\right) = \exp\left(-\left(\frac{60}{51.67}\right)^{2.649}\right) \approx \exp(-1.485) \approx 0.2264.$$

(c) Sea $Z \sim$ Logística(μ, σ). Del formulario tenemos que $\mathbb{E}(Z) = \mu$ y $\sqrt{\mathrm{Var}(Z)} = \sigma\pi/\sqrt{3} = 7$. Así, $\mu = 15$ y $\sigma = 7\sqrt{3}/\pi = 3.859$.

Se pide $P(Z < 5) = F(5)$, donde $F(x)$ es dado en formulario:

$$P(Z < 5) = F(5) = \frac{\exp\left((5-15)/3.859\right)}{1 + \exp\left((5-15)/3.859\right)} = \frac{0.0749}{1 + 0.0749} \approx 0.0696$$

Asignación de Puntaje:

a)

Logro 1 [2 ptos]: por obtener los parámetros de la distribución Log-Normal y calcular lo solicitado.

- [1 ptos] por obtener $\lambda = \log(12)$ y $\xi = 0.429$.
- [1 ptos] por calcular $P(X > 24) = 0.053$.

b)

Logro 2 [2 ptos]: por obtener los parámetros de la distribución Weibull y calcular lo solicitado.

- [1 ptos] por calcular $\beta = 2.649$ y $\eta = 51.67$.
- [1 ptos] por obtener $P(Y > 60) = 0.2264$.

c)

Logro 3 [2 ptos]: por obtener los parámetros de la distribución Logística y calcular lo solicitado.

- [1 pto] por plantear la expresión de la función de distribución logística.
- [1 pto] por calcular $P(Z < 5) = 0.0696$.

Pregunta 4

La empresa ChipIA se dedica a vender chips que incorporan métodos de inteligencia artificial, destinados a ser integrados en dispositivos electrónicos fabricados por terceros. Durante la fase inicial, los chips aún pueden presentar fallas, por lo que fueron ofrecidos a la microempresa TelefonIA para pruebas.

(a) **[2.0 Ptos.]** TelefonIA probó 10 chips, cada uno integrado de manera independiente en uno de dos dispositivos: A o B. La probabilidad de que un chip sea usado en un dispositivo A es de 60 %, y en un dispositivo B, de 40 %. Si los chips instalados en A fallan con probabilidad 7 %, y en B con probabilidad 9 %, ¿cuál es la probabilidad de que, entre los 10 chips, fallen exactamente 3?

(b) **[2.0 Ptos.]** Tras la prueba anterior, TelefonIA compró una caja con 30 chips, pero la devolvió pues al tomar una muestra aleatoria de 4 de estos, encontraron 3 chips defectuosos. Al revisar la caja, ChipIA confirmó que contenía exactamente 4 chips defectuosos. ¿Cuál es la probabilidad de que ocurriera lo observado por la microempresa?

(c) **[2.0 Ptos.]** A raíz del incidente, ChipIA revisó su inventario, donde se estima que un 1 % de los chips presenta fallas. ¿Cuántos chips se espera revisar hasta encontrar 4 defectuosos?

Solución

(a) Sea X la cantidad de chips fallados entre los 10 probados. Como X representa número de éxitos de un total de 10, $X \sim$ Binomial$(10, p)$, donde p es la probabilidad de falla de un chip. Como esta falla depende del tipo de dispositivo en el que fue instalado, aplicamos el teorema de probabilidad total para obtener p:

$$p = P(\text{Falla}) = P(A) \cdot P(\text{Falla} \mid A) + P(B) \cdot P(\text{falla} \mid B) = 0.6 \cdot 0.07 + 0.4 \cdot 0.09 = 0.042 + 0.036 = 0.078$$

Buscamos $P(X = 3)$:

$$P(X = 3) = \binom{10}{3} (0.078)^3 (1 - 0.078)^7 \approx 120 \cdot (0.000475) \cdot (0.5445) \approx 0.032$$

(b) Este caso sigue un modelo hipergeométrico: de una población de 30 chips con 4 defectuosos, se extraen 4 sin reemplazo y se observan 3 defectuosos. Sea X la cantidad de defectuosos en la muestra, $X \sim$ Hipergeom$(N = 30, m = 4, n = 4)$. Buscamos $P(X = 3)$:

$$P(X = 3) = \frac{\binom{4}{3}\binom{26}{1}}{\binom{30}{4}} = \frac{4 \cdot 26}{27405} = \frac{104}{27405} \approx 0.0038$$

(c) Sea N el número de chips a revisar hasta encontrar 4 defectuosos, el cual sigue una distribución Binomial Negativa con parámetros $r = 4$, $p = 0.01$. Así, lo pedido es:

$$\mathbb{E}(N) = \frac{r}{p} = \frac{4}{0.01} = 400$$

Asignación de Puntaje:

a)

Logro 1 [1 ptos]: por identificar modelo Binomial y sus parámetros.

- [0.5 ptos] por indicar distribución Binomial.
- [0.5 ptos] por indicar $n = 10$ y que p es desconocido y debe calcularse.

Logro 2 [1 ptos]: por calcular p y probabilidad solicitada.

- [0.5 ptos] por obtener $p = 0.078$ mediante probabilidades totales.
- [0.5 ptos] por obtener $P(X = 3) = 0.032$.

b)

Logro 3 [1.5 ptos]: por identificar correctamente el modelo hipergeométrico y sus parámetros.

- [0.5 ptos] por identificar modelo hipergeométrico.
- [1 pto] por obtener $N = 30$ [0.3 pto], $m = 4$ [0.4 pto] y $n = 4$ [0.3 pto].

Logro 4 [0.5 ptos]: por obtener probabilidad solicitada, $P(X = 3) = 0.0038$.

c)

Logro 5 [1.5 ptos]: por identificar correctamente el modelo binomial negativo y sus parámetros.

- [0.5 pto] por reconocer que el número de chips hasta encontrar 4 defectuosos es una distribución Binomial Negativa.
- [1 pto] por indicar que $r = 4$ [0.5 pto] y $p = 0.01$ [0.5 pto].

Logro 6 [0.5 ptos]: por obtener valor esperado, $\mathbb{E}(N) = 400$.