Universidad de la República | **Probabilidad y Estadística**
Facultad de Ingeniería - IMERL | **Segundo Semestre 2016**

Examen Diciembre
Lunes 19 de diciembre de 2016.

Número de Parcial	Cédula	Nombre y Apellido

PARA USO DOCENTE			
Ej. 1	Ej. 2	Ej.3	TOTAL

Ejercicio 1. [**20 puntos**]

Cierto entrenador de baby fútbol tiene predilección por los niños zurdos. Si un niño es zurdo, juega en todas las fechas del campeonato. Si es diestro, tiene probabilidad 2/3 de jugar en cualquier fecha dada, y no depende de cuántas veces haya jugado antes. Sabemos que 1/10 de los niños del club son zurdos.

1. (a) Si Pedro juega en la primera fecha, ¿cuál es la probabilidad de que sea zurdo?
   (b) Si Pedro juega en las primeras tres fechas ¿cuál es la probabilidad de que sea zurdo?
2. En este club, cada niño tiene una camiseta con su nombre en la espalda. Luis y Celso son dos mellizos que el entrenador es incapaz de distinguir. Luis es zurdo y Celso es diestro. Cada fin de semana estos dos niños tiran una moneda equilibrada, y si sale cara se intercambian las camisetas, engañando al entrenador.
   (a) ¿Cuál es la probabilidad de que Celso juege un determinado partido? ¿Y Luis?
   (b) Si el campeonato tiene doce fechas, ¿cuántas veces esperamos que juegue Luis? Justifique su respuesta.

Ejercicio 1. Solución
Se consideran los sucesos:

- $P_1 = \{\text{Pedro juega en la primer fecha}\}$
- $Z = \{\text{el jugador es zurdo}\}$

1. Por la fórmula de Bayes tenemos que
$$\mathbb{P}(Z|P_1) = \frac{\mathbb{P}(P_1|Z)\,\mathbb{P}(Z)}{\mathbb{P}(P_1)} = \frac{1 \times 1/10}{\mathbb{P}(P_1)}.$$
Además por la fórmula de la probabilidad total se tiene que:
$$\mathbb{P}(P_1) = \mathbb{P}(P_1|Z)\,\mathbb{P}(Z) + \mathbb{P}(P_1|Z^c)\,\mathbb{P}(Z^c) = 1 \times \frac{1}{10} + \frac{2}{3} \times \frac{9}{10} = \frac{7}{10}.$$
Por lo tanto, la probabilidad es $\mathbb{P}(Z|P_1) = \frac{1}{7}$.
2. En esta segunda parte tenemos que usar la propiedad de que jugar en diferentes fechas son sucesos independientes. Sea P_3 el suceso $P_3 = \{\text{Pedro juega en las primeras tres fechas}\}$. Entonces tenemos que:

- $\mathbb{P}(P_3|Z) = \mathbb{P}(P_1|Z)^3 = 1$ y $\mathbb{P}(P_3|Z^c) = \mathbb{P}(P_1|Z^c)^3 = (2/3)^3 = \frac{8}{27}$.

Por lo tanto:

$$\mathbb{P}(P_3) = \mathbb{P}(P_3|Z)\,\mathbb{P}(Z) + \mathbb{P}(P_3|Z^c)\,\mathbb{P}(Z^c) = \frac{1}{10} + \frac{8}{27}\frac{9}{10} = \frac{11}{30}.$$

De donde

$$\mathbb{P}(Z|P_3) = \frac{\mathbb{P}(P_3|Z)\,\mathbb{P}(Z)}{\mathbb{P}(P_3)} = \frac{3}{11}.$$

3. Consideremos ahora los sucesos:
   - $C = \{\text{Celso juega}\}$
   - $L = \{\text{Luis juega}\}$
   - $I = \{\text{Celsos y Luis intercambiaron camisetas}\}$.

$$\mathbb{P}(C) = \mathbb{P}(C|I)\,\mathbb{P}(I) + \mathbb{P}(C|I^c)\,\mathbb{P}(I^c) = 1\frac{1}{2} + \frac{2}{3}\frac{1}{2} = \frac{5}{6}.$$

La probabilidad de que Luis juegue es igual a la probabilidad de que Celso juegue.

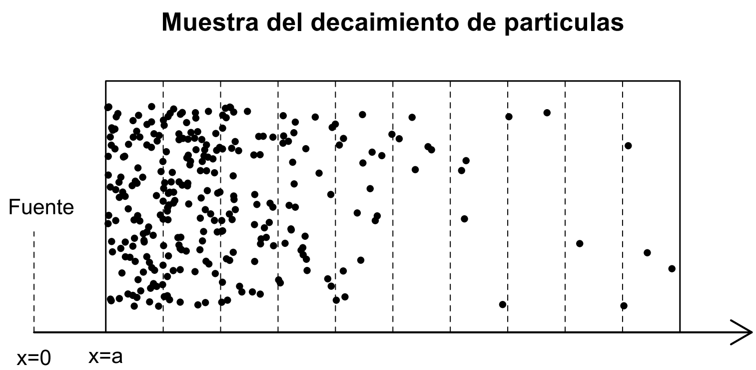
4. Sea X una variable aleatoria que cuenta el número de partidos del campeonato en que Luis juega. Dado que jugar en diferentes fechas son sucesos independientes y de que la probabilidad de jugar es siempre igual, resulta que X tiene distribución Binomial de parámetros $n = 12$ y $p = 5/6$. Por lo tanto en número esperado de partidos en los que juega Luis está dado por:

$$\mathrm{E}(X) = np = 12\frac{5}{6} = 10.$$

Ejercicio 2. [**45 puntos**]

Partículas inestables se emiten desde una fuente ubicada en la posición $x = 0$ como en la figura. Cada partícula viaja en línea recta horizontalmente y se desintegra a una distancia X de la fuente. Se asume que X tiene distribución exponencial de parámetro desconocido $\lambda > 0$.

Sin embargo, estos eventos de desintegración se pueden observar solamente si ocurren en una ventana $V = [a, \infty)$. Llamemos X_V a la distancia de desintegración de las partículas observadas en V. En la siguiente figura se muestran las partículas que se desintegran en la ventana V. La porción de la ventana donde se oberservan partículas desintegradas se ha dividido en 10 intervalos de longitud 10 cm, y se muestra la cantidad de partículas desintegradas en cada intervalo. El total de partículas en la ventana es 287.



Frec = 91 88 48 29 15 5 4 2 1 4

Total = 287

1. (a) Hallar $p(\lambda, a)$ la probabilidad de que una partícula se desintegre fuera de la ventana V.
   (b) Suponiendo que la cantidad de partículas emitidas por la fuente es $N = 500$ y que $a = 10$ cm, hallar una estimación para λ.

2. En esta parte suponemos que se desconoce la cantidad N de partículas emitidas por la fuente.

   (a) Probar que para todo $x \geq a$, la probabilidad condicional de $\{X \leq x\}$ dado que la partícula se desintegra en la ventana $\{X \in V\}$ está dada por
   $$P(X \leq x | X \in V) = 1 - e^{-\lambda(x-a)}.$$

   (b) Denotamos por $F_V(x) = P(X \leq x | X \in V)$, $x \in \mathbb{R}$ a la función de distribución de X_V. Hallar la densidad y el valor esperado de X_V.

   (c) i. Sea $X_V^1, \ldots, X_V^n$ una muestra i.i.d. de distribución F_V. Hallar el estimador de máxima verosimilitud λ_{MV} de λ.

   ii. Sabiendo que $a = 10$ y $\sum_{i=1}^{i=287} X_V^i = 7399.711$, calcular λ_{MV}.

   iii. Con los mismos datos que antes y asumiendo que $\lambda = \lambda_{MV}$, dar una estimación para la cantidad N de partículas emitidas.

Ejercicio 2. Solución

1. (a) $p(\lambda, a) = \mathbb{P}(X \in [0, a]) = F_X(a) = 1 - e^{-\lambda a}$.

   (b) Sabemos que de $N = 500$ partículas emitidas, 287 se desintegraron en la ventana. Por lo tanto 213 se desintegraron fuera de la venta, de donde podemos aproximar $p(\lambda, a) = 1 - e^{-\lambda a} \approx \frac{213}{500}$. De donde, usando que $a = 10$ resulta que $\lambda \approx 0.055$.

2. (a) Observar que si $x < a$ entonces $P(X \leq x | X \in V) = 0$. Para $x \geq a$:
   $$\mathbb{P}(X \leq x | X \in V) = \frac{\mathbb{P}(X \leq x, X \in V)}{\mathbb{P}(X \in V)} = \frac{\mathbb{P}(X \in [a, x])}{1 - (1 - e^{-\lambda a})}$$
   $$= \frac{-e^{-\lambda x} + e^{-\lambda a}}{e^{-\lambda a}} = 1 - e^{-\lambda(x-a)}.$$

   (b) Tenemos que para $x \geq a$, $F_{X_V}(x) = 1 - e^{-\lambda(x-a)}$. Por lo tanto, la densidad está dada por $f_{X_V}(x) = F'_{X_V}(x)$, esto es:
   $$f_{X_V}(x) = \begin{cases} \lambda e^{-\lambda(x-a)} & x \geq a \\ 0 & x < a \end{cases}$$
   Dado que X_V es positiva, podemos calcular su valor esperado a partir de la función de distribución;
   $$\mathrm{E}(X_V) = \int_0^\infty 1 - F_{X_V}(x) dx = \int_0^a 1 dx + \int_a^\infty e^{-\lambda(x-a)} dx = a + \frac{1}{\lambda}.$$
   Se llega al mismo resultado a partir de la fórmula del valor esperado con la densidad y utilizando integración por partes.

   (c) i. Consideremos la función de verosimilitud:
   $$\mathcal{L}(\lambda) = \prod_{i=1}^n f_{X_V}(X_V^i) = \prod_{i=1}^n \lambda e^{-\lambda(X_V^i - a)} = \lambda^n e^{n\lambda a} e^{-\lambda \sum_{i=1}^n X_V^i}$$
   Tomando logaritmo tenemos:
   $$\ln(\mathcal{L}(\lambda)) = n \ln(\lambda) + n\lambda a - \lambda \sum_{i=1}^n X_V^i$$
   Y derivando respecto a λ para obtener los extremos resulta que:
   $$\frac{\partial \ln(\mathcal{L}(\lambda))}{\partial \lambda} = \frac{n}{\lambda} + na - \sum_{i=1}^n X_V^i = 0$$

de donde

$$\hat{\lambda}_{MV} = \frac{1}{\frac{\sum_{i=1}^{n} X_V^i}{n} - a} = \frac{1}{\overline{X}_V - a}$$

Es fácil ver que es un máximo y que además coincide con el estimador por momentos de λ.

ii. Sabemos que $a = 10$ y que $\sum_{i=1}^{i=287} X_V^i = 7399.711$, de donde $\overline{X}_V = \frac{7399.711}{287} = 25.78$. Por lo tanto

$$\hat{\lambda}_{MV} = \frac{1}{25.78 - 10} = 0.063$$

.

iii. Por la parte 1 del ejercicio sabemos que para $\lambda = \hat{\lambda}_{MV}$,

$$p(\lambda, a) = 1 - e^{-10*0.063} = 0.467 \approx \frac{N - 287}{N} = 1 - \frac{287}{N}$$

de donde despejando obtenemos que es posible estimar N por $\hat{N} = \frac{287}{1-0.467} = 538$ (o 539) partículas.

Ejercicio 3. [**35 puntos**]

El número de autos N que pasan por un peaje del Uruguay en un intervalo de 1 minuto, se puede describir mediante una variable aleatoria Poisson de parámetro λ. La siguiente tabla muestra los datos de la cantidad de autos contabilizados en 100 días durante el mismo minuto cada día:

Nro. de autos	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Frecuencia	2	1	2	7	7	12	14	17	8	13	5	6	3	3

1. Realizar un histograma de N. Indicar, primer cuartil, mediana, moda y tercer cuartil de los datos.
2. A partir de los datos, estimar la esperanza y la varianza de N.
3. Hallar un intervalo de confianza aproximado al nivel 0.9 para λ.
4. (a) Plantear un test de hipótesis aproximado para testear:

$$\begin{cases} H_0 : \lambda = 10 \\ H_1 : \lambda < 10 \end{cases}$$

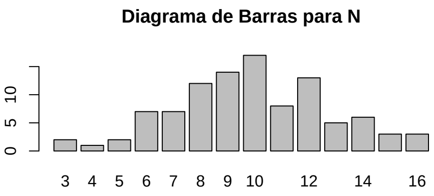
   (b) ¿Cuál es la decisión para los niveles $\alpha = 0.01$, $\alpha = 0.1$ y $\alpha = 0.5$?
   (c) Hallar el p-valor.
   (d) Para la hipótesis alternativa $H_1 : \lambda = 9.99$ y $\alpha = 0.1$, hallar la probabilidad de error de tipo II.

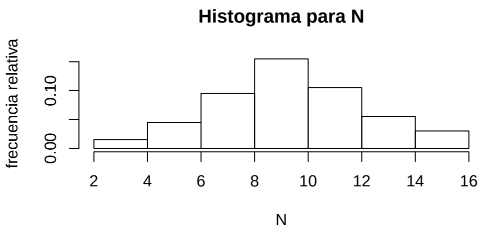
Ejercicio 3. Solución

1. Primer y tercer cuartil : $Q_1 = 8$ y $Q_3 = 12$. Mediana $m_N = 10$ y moda $m = 10$.
2. Por la ley de los grandes números, la esperanza la aproximamos por el promedio, esto es $\overline{N}_{100} = 9.89$. Para estimar la varianza podemos utilizar el estimador $s_{100}^2 = 8.12$. También es posible argumentar que por ser los datos Poisson, entonces $\mathrm{Var}(N) = \lambda$ y por lo tanto, también es posible estimar la varianza con el promedio.
3. Un intervalo de confianza aproximado para $\lambda = \mathrm{E}(N)$ al nivel de confianza $1 - \alpha$ está dado por:

$$I_\alpha = \left[\overline{N}_{100} - \frac{s_{100} z_{\alpha/2}}{10}, \overline{N}_{100} + \frac{s_{100} z_{\alpha/2}}{10}\right]$$

Para $\alpha = 0.9$, tenemos que $z_{\alpha/2} = 1.65$ y el intervalo de confianza resulta: $I_\alpha = [9.42, 10.36]$. Si se utiliza el promedio como estimador de la varianza el intervalo resulta: $I_\alpha = [9.37, 10.41]$.





4. (a) La región crítica al nivel α para un test de hipótesis para la media basado en el TCL para datos Poisson está dada por

$$\mathcal{R}_\alpha = \left\{ \sqrt{100}\frac{\overline{N}_{100} - \lambda_0}{\sqrt{\lambda_0}} \leq -z_\alpha \right\}$$

(b) Para los datos que tenemos resulta que $\mathcal{R}_\alpha = \{-0.347 \leq -z_\alpha\}$. Por lo tanto,

- si $\alpha = 0.01$, $z_{0.01} = 2.33$ y No se rechaza H_0.
- si $\alpha = 0.1$, $z_{0.1} = 1.28$ y No se rechaza H_0.
- si $\alpha = 0.5$, $z_{0.5} = 0$ y Se rechaza H_0.

(c) El p-valor está dado por $\alpha^* = \Phi(-0.347) = 0.375$.

(d) Para $\alpha = 0.1$, la región crítica es $\mathcal{R}_{0.1} = \left\{ \sqrt{100}\frac{(\overline{N}_{100} - 10)}{\sqrt{10}} \leq -1.28 \right\}$. Por lo tanto,

$$\beta = P_{H_1}(\mathcal{R}^c_{0.1}) = P_{H_1}(\overline{N}_{100} - 10 > -0.41) = P_{H_1}(\overline{N}_{100} > 9.59) =$$
$$P_{H_1}\left(\sqrt{100}\frac{(\overline{N}_{100} - 9.99)}{\sqrt{9.99}} > \sqrt{100}\frac{(9.59 - 9.99)}{\sqrt{9.99}} \right) \approx 1 - \Phi(-1.26) = 0.82.$$

Donde en el último paso usamos que bajo H_1 la verdadera media y varianza es $\lambda_1 = 9.99$.

Obs: Se considerará como correcto la región crítica definida con s_n como estimador del desvío estándar..