Clase Auxiliar 3 Probabilidades y Estadística. Prim '08 FC, VR

*) Variables Aleatorias Discretas

1°) Aproximación de una variable aleatoria con ley Poisson.

Def Se dice que X tiene ley Poisson (/Puaˀson/) de parametro λ si su densidad está dada por

$$f_X(i) = e^{-\lambda}\frac{\lambda^i}{i!} \qquad \text{si } i=0,1,\ldots$$

Se interpreta como el # de llegadas de individuos a algún proveedor de servicios, por unidad de tiempo. La tasa λ es en unidades $\left[\frac{\text{units}}{\text{u tiempo}}\right]$, y representa una tasa promedio ("en promedio llegan 20 personas por hora").

La ley de Poisson puede ser vista como una aproximación de variables aleatorias con ley Geom(n,p) donde n y p son tales que 'np' es de tamaño "moderado".

Para ver lo anterior, supongamos que $X \sim \text{Binom}(n,p)$ y sea $\lambda = np$. Entonces

$$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i} = \frac{n!}{(n-i)!\,i!}\, p^i (1-p)^{n-i}$$

$$= \frac{n!}{(n-i)!\,i!}\left(\frac{\lambda}{n}\right)^i \left(1-\frac{\lambda}{n}\right)^{n-i}$$

$$= \underbrace{\frac{n(n-1)\cdots(n-(i-1))}{n^i}}_{\downarrow\, 1} \cdot \frac{1}{i!}\,\lambda^i \underbrace{\left(1-\frac{\lambda}{n}\right)^n}_{\downarrow\, e^{-\lambda}} \underbrace{\left(1-\frac{\lambda}{n}\right)^{-i}}_{\downarrow\, 1}$$

$$\xrightarrow{n\to\infty} \frac{e^{-\lambda}\lambda^i}{i!}$$

Recordar entonces que la distribución de Poisson está asociada a números de la forma:

"el # de "[algo]" que ocurre en cierto periodo de tiempo"

Ejemplos: › "El número de personas que llega a la fila de una caja de supermercado durante 15 minutos"

› El número de micros que llega a un paradero durante una hora.

› Etc.

Problema] Suponga que el # de errores tipográficos en cada página de un libro tiene una distribución de Poisson con parámetro $\lambda = 1/2$.

a) Calcule la probabilidad de que haya al menos un error en cierta página (dada).

Sol Sea X la v.a. que representa el # de errores tipográficos en dicha página.

Entonces $X \sim \text{Poisson}(\lambda)$. Nos piden $\mathbb{P}(X \geq 1)$

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X = 0)$$

$$= 1 - \frac{e^{-\lambda}\lambda^0}{0!} = 1 - e^{-\lambda}$$

$$= 1 - e^{-1/2} \sim 0{,}393.$$

b) Para el caso anterior, calcule la probabilidad de que el # de errores en dicha página es par.

Sol Piden $\mathbb{P}(X \text{ es par}) = \mathbb{P}(\{X=0\} \cup \{X=2\} \cup \{X=4\} \cup \dots \cup \{X=2k\} \cup \dots)$

$$= \sum_{k=0}^{\infty} \mathbb{P}(X = 2k)$$

$$= \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^{2k}}{(2k)!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!}$$

$$\boxed{\cosh(x) = \frac{e^x + e^{-x}}{2}}$$

$$= e^{-\lambda}\cosh(\lambda) = e^{-\lambda}\left(\frac{e^{\lambda} + e^{-\lambda}}{2}\right) = \frac{e^{-2\lambda} + 1}{2}$$

2º) Ley Geometrica.

Modela la primera vez que se obtiene un resultado favorable después de realizar repetidamente un experimento y obtener resultados desfavorables (realizaciones de experimentos independientes)

Sea X v.a. se dice que $X \sim Geom(p)$ si su densidad es

$$f_X(i) = (1-p)^{i-1} p \quad , \quad i=1,2,\ldots$$

Notemos que si p es la probabilidad de exito y $(1-p)$ la probabilidad de fracaso,

$$(1-p)^{i-1}p = \underbrace{(1-p)}_{\substack{\text{fracaso} \\ \text{intento 1}}}\underbrace{(1-p)}_{\substack{\text{fracaso} \\ \text{intento 2}}}\cdots\underbrace{(1-p)}_{\substack{\text{fracaso} \\ \text{intento } i-1}}\underbrace{p}_{\substack{\text{exito} \\ \text{intento } i}}$$

3º) Ley Binomial

Modela la elección de k objetos entre n, cada uno con probabilidad p de ser elegido e independientemente de los demas.

Sea $X \sim Binom(n,p)$. Entonces la densidad de X es

$$f_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k=0,1,\ldots,n$$

Problema 2 Una urna contiene N bolas blancas y n bolas negras. Se extraen bolas al azar, una cada instante, hasta que una bola negra es sacada. Si cada bola es repuesta cada vez que se extrae y cada extracción es independiente de las otras,

a) ¿Cual es la probabilidad de que se necesiten hacer exactamente n extracciones?

b) ¿Cual es la probabilidad de que se necesiten al menos k extracciones?

sol Sea X la v.a. que indica el # de extracciones necesarias hasta que se saca una bola negra. Es claro que el soporte de esta v.a. es $\{1,2,3,\ldots\}$.

Ahora, veamos la densidad de X: Sean G_i el evento "una bola negra sale por 1a vez en el intento i" y S_i el evento "en el intento i salió una bola negra".

Entonces $G_i = S_1^c \cap S_2^c \cap \dots \cap S_{i-1}^c \cap S_i$

$$
\begin{aligned}
y \quad \mathbb{P}(G_i) &= \mathbb{P}(S_1^c \cap S_2^c \cap S_3^c \cap \dots \cap S_{i-1}^c \cap S_i) \\
&= \mathbb{P}(S_1^c) \cdot \mathbb{P}(S_2^c) \cdot \dots \cdot \mathbb{P}(S_{i-1}^c)\,\mathbb{P}(S_i) \qquad \text{(independencia de las extracciones)} \\
&= \underbrace{(1-p)(1-p)\cdots(1-p)}_{i-1 \text{ veces}}\, p \\
&= (1-p)^{i-1} p
\end{aligned}
$$

¿Quién es p? $p = \mathbb{P}(S_1) =$ Prob. de sacar en la extracción 1 una bola negra.

$$= \frac{n}{n+N}$$

$$\Rightarrow \mathbb{P}(X=i) = \mathbb{P}(G_i) = (1-p)^{i-1} p$$

$$\Rightarrow X \sim \text{Geom}(p)$$

Ahora,

a) Hacer exactamente n extracciones significa que la 1a bola negra sale en la ocasión n. Esto es $\{X=n\}$.

$$\mathbb{P}(X=n) = (1-p)^{n-1} p$$

b) Necesitar al menos k extracciones es el evento $\bigcup_{n\geq k} \{X=n\} = \{X \geq k\}$

$$\Rightarrow \mathbb{P}(X\geq k) = \sum_{n\geq k} \mathbb{P}(X=n) = \sum_{n=k}^{\infty} (1-p)^{n-1} p = p(1-p)^{k-1} \sum_{n=0}^{\infty} (1-p)^n$$

$$= p(1-p)^{k-1} \frac{1}{\cancel{1}+\cancel{(1}-p)} = (1-p)^{k-1}$$

i) (Propuesto) Hacer lo mismo, pero en el caso que el experimento (en cada instante) sea:

"Elegir k bolas de entre las $M+N$ de la urna, y parar la 1ª vez que se tienen n bolas negras de esas k."

¿como cambia p? ¿Cambian los eventos de interés?

Problema 3 Sean $\mathbb{C}_i \sim Geom(p_i)$ variables aleatorias independientes.

Demostrar que la variable aleatoria $\mathbb{C} = \min_{i=1,\dots,m} \mathbb{C}_i$ se distribuye Geometrica.

Dem Veamos la ley de una variable geometrica:

$$\mathbb{P}(\mathbb{C}_i \geq n) = \sum_{k=n}^{\infty} \mathbb{P}(\mathbb{C}_i = k) = \sum_{k=n}^{\infty} (1-p_i)^{k-1} p_i$$

$$= p_i (1-p_i)^{n-1} \sum_{k=0}^{\infty} (1-p_i)^k = \cancel{p_i} (1-p_i)^{n-1} \frac{1}{\cancel{1}-\cancel{(1}-p_i)}$$

$$= (1-p_i)^{n-1}$$

Nos basta probar entonces que $\mathbb{P}(\mathbb{C} \geq n) = (1-p)^{n-1}$ algún $p \in (0,1)$

Ahora, $n \geq 1$

$$\mathbb{P}(\mathbb{C} \geq n) = \mathbb{P}\left(\min_{i=1,\dots,m} \mathbb{C}_i \geq n\right) = \mathbb{P}(\mathbb{C}_1 \geq n, \dots, \mathbb{C}_m \geq n)$$

$$= \mathbb{P}(\mathbb{C}_1 \geq n)\, \mathbb{P}(\mathbb{C}_2 \geq n) \cdots \mathbb{P}(\mathbb{C}_m \geq n) \qquad \text{(por indep)}$$

$$= (1-p_1)^{n-1} \cdots (1-p_m)^{n-1} = \Big[\underbrace{\underbrace{(1-p_1)}_{\leq 1} \cdots \underbrace{(1-p_m)}_{\leq 1}}_{\leq 1}\Big]^{n-1}$$

$$= \Big(1 - \underbrace{\{1-(1-p_1)\cdots(1-p_m)\}}_{p \in (0,1)}\Big)^{n-1}$$

$\Rightarrow \mathbb{C} \sim Geom(p)$.

Obs En el caso $m=2$, la interpretación es que se están echando a correr dos "tiempos" discretos. La distribución del menor de esos dos tiempos es tb una geométrica, con parámetro $p = 1-(1-p_1)(1-p_2) = 1-(1-p_1-p_2+p_1p_2) = p_1+p_2-p_1p_2$.

Problema 4 En el caso anterior, con $m=2$, ¿cuál es la probabilidad de que e_1 ocurra antes que e_2?

Sol Se pide $\mathbb{P}(e_1 < e_2)$. Para ello, usamos probabilidades condicionales

$$\mathbb{P}(e_1<e_2) = \sum_{k=1}^{\infty} \mathbb{P}(e_1<e_2 \mid e_1=k)\,\mathbb{P}(e_1=k)$$

$$= \sum_{k=1}^{\infty} \mathbb{P}(k<e_2 \mid e_1=k)\,\mathbb{P}(e_1=k)$$

$$= \sum_{k=1}^{\infty} \mathbb{P}(k<e_2)\,\mathbb{P}(e_1=k) \qquad \text{independencia de } e_1 \text{ y } e_2$$

$$= \sum_{k=1}^{\infty} (1-p_2)^k (1-p_1)^{k-1} p_1$$

$$= p_1(1-p_2) \sum_{k=0}^{\infty} [(1-p_1)(1-p_2)]^k$$

$$= p_1(1-p_2) \cdot \frac{1}{1-(1-p_1)(1-p_2)}$$

$$= \frac{p_1(1-p_2)}{p}$$

Problemas Dos estudiantes compiten por un premio. La competencia consiste en D preguntas. Se hace una pregunta y el estudiante que no sepa resolverla queda eliminado inmediatamente. Un estudiante i tiene una probabilidad q_i en no contestar bien alguna de las preguntas. Si un estudiante i logra terminar toda la prueba, el tiempo que transcurre para terminarla se distribuye Geometrica de parametro λ_i. Gana el estudiante que termina la prueba en menor tiempo.

a) ¿Cual es la probabilidad de que un estudiante i termine la prueba? cual es la probabilidad de que conteste al menos k preguntas bien?

Sol Si consideramos X_i la va que representa el instante de fallo del jugador i, vemos que X_i puede tomar los valores $1, 2, \dots, D$, y $D+1$ si no falla la pregunta D.

Entonces que el estudiante i termine la prueba es equivalente a que $\{X_i = D+1\}$.
Esto significa que el estudiante no falló ninguna pregunta.

$$\Rightarrow \mathbb{P}(X_i = D+1) = (1-q_i)^D = \underbrace{(1-q_1)}_{\text{prob de contestar bien pregunta 1}} \cdots \underbrace{(1-q_D)}_{\text{prob de contestar bien pregunta D.}}$$

En el otro caso, la probabilidad de contestar bien al menos k preguntas bien es igual a

$$\mathbb{P}(X_i \geq k+1) = \sum_{j=k+1}^{D+1} \mathbb{P}(X_i = j) = \sum_{j=k+1}^{D} (1-q_i)^{j-1} q_i + (1-q_i)^D$$

$$= (1-q_i)^k$$

b) Cual es la probabilidad, dado que ambos estudiantes terminaron la prueba, de que el estudiante 1 sea el ganador?

Sol Las leyes de los tiempos condicional a que terminaron la prueba son $Geom(\lambda_i)$. Por el problema anterior, la probabilidad pedida es (si e_i es el tiempo del jugador i)

$$\mathbb{P}\left(e_1 < e_2 \,\middle|\, \begin{array}{c}\text{ambos terminaron}\\ \text{la prueba}\end{array}\right) = \frac{\lambda_1 (1-\lambda_2)}{\lambda} \qquad \text{con } \lambda = 1-(1-\lambda_1)(1-\lambda_2).$$

c) ¿Cuál es la probabilidad de que el estudiante 1 sea el ganador?

Sol Sean los eventos G_1 = "Estudiante 1 gana", T_i = "Estudiante i termina la prueba".

Entonces $P(G_1) = P(G_1 | T_1) P(T_1) + \overset{\to 0}{\cancel{P(G_1 | T_1^c)}} P(T_1^c)$

$$= P(G_1 \cap T_1) = P(G_1 \cap T_1 \cap T_2) + P(G_1 \cap T_1 \cap T_2^c)$$

$$= P(G_1 | T_1 \cap T_2) P(T_1 \cap T_2) + P\underset{\to 1}{(\cancel{G_1 / T_1 \cap T_2^c})} P(T_1 \cap T_2^c)$$

$$= P(G_1 / T_1 \cap T_2) P(T_1) P(T_2) + P(T_1) P(T_2^c)$$ (independencia entre T_1 y T_2)

$$= \frac{\lambda_1 (1-\lambda_2)}{\lambda} (1-q_1)^D (1-q_2)^D + (1-q_1)^D (1-(1-q_2)^D)$$