Tarea 4 - Análisis de datos
Giovanni Gamaliel López Padilla

Problema 3

Completa cada linea con $\leq, \geq, =$ para obtener una verdad para cualquier v.a. discreta (X,Y)

- **Entropia_Shannon(X+1) ... Entropia_Shannon(X)**

  Sea $H(X+1) = H(X+1, X)$, entonces:

$$\begin{aligned} H(X+1,X) &= Elog_2\left(\frac{1}{P(X+1,X)}\right) \\ &= Elog_2\left(\frac{1}{P(X+1|X)P(X)}\right) \\ &= E\left(-log_2\left(P(X+1|X)P(X)\right)\right) \\ &= E\left(-log_2(P(X+1|X)) - log_2(P(X))\right) \end{aligned}$$

  por linealidad de la esperanza se obtiene que:

$$\begin{aligned} H(X+1,X) &= E\left(-log_2(P(X+1|X)) - log_2(P(X))\right) \\ &= E(-log_2(P(X+1|X))) + E(-log_2(PX)) \\ &= H(X+1|X) + H(X) \end{aligned}$$

  como H¿0 entonces:

$$H(X+1,X) > H(X)$$

  por lo tanto:

$$H(X+1) > H(X)$$

- **Var(X+1) ... Var(X)**

  Desarrollando $Var(X+1)$ se tiene lo siguiente:

$$\begin{aligned} Var(X+1) &= E(X+1)^2 - (E(X+1))^2 \\ &= E(X^2+2X+1) - ((EX)^2 + 2EX + 1) \end{aligned}$$

  Aplicando la linealidad de la esperanza de tiene que

$$\begin{aligned} Var(X+1) &= E(X^2+2X+1) - ((EX)^2 + 2EX + 1) \\ &= EX^2 + E(2X) + E(1) - (EX)^2 - 2EX - 1 \end{aligned}$$

Aplicando $E(a) = a$ y $E(aX) = aEX$, para $a \in \mathbb{R}$, se obtiene que:

$$\begin{aligned} Var(X+1) &= EX^2 + E(2X) + E(1) - (EX)^2 - 2EX - 1 \\ &= EX^2 + 2EX + 1 - (EX)^2 - 2EX - 1 \\ &= EX^2 - (EX)^2 \\ &= Var(X) \end{aligned}$$

por lo tanto:

$$Var(X+1) = Var(X)$$

- **EX ... E exp(X)**

  Expandiendo hasta el segundo termino de la serie de potencias de EX se obtiene lo siguiente:

  $$exp(x) > 1 + x$$

  entonces

  $$E(exp(X)) > E(1 + X)$$

  por linealidad de la esperanza entonces:

  $$\begin{aligned} E(exp(X)) &> E(1 + X) \\ &> E(1) + E(X) \\ E(exp(X)) &> 1 + E(X) \end{aligned}$$

  como

  $$E(X) + 1 > EX$$

  podemos afirmar que

  $$E(X) < E(exp(X))$$

Problema 4

Problema 4a

Supongamos que te contrataron como científico de datos en una base aerea durante la segunda guerra mundial. Cada dia salen aviones para misiones de bombardeo. Estudias cada noche el daño en los aviones que regresan de su misión. Aplicas tu algoritmo favorito de aprendizaje máquina y reconocimiento de patrones para buscar cuáles partes se dañan más y por ende valen la pena reforzar en los aviones. ¿Qué sera una limitante muy fuerte para poder llegar a conclusiones contundentes?
La limitante que tendria sería el no conocer los daños que sufrieron los aviones que no regresan.

Problema 4b

Quieres estimar la duración promedio de la estancia de turistas en la ciudad de Guanajuato. Eliges en diferentes momentos, en diferentes lugares un turistas al azar y preguntas por la duración de su estancia. Decides promediar los valores obtenidos para estimar la duración promedio. ¿Te parece una buena estrategía?
No me parece una buena estrategía ya que pueden existir diferncias grandes en el tiempo de estancia en periodos vacacionales. Esto se podria comprobar calculando las variancas. Para obtener una mejor aproximación sería agregar una variable temporal al promedio, esto para que se pueda introducir la variación existente en periodos vacacionales.

Problema 5

Un fabricante de misiles pretende que la precisión de sus misiles de larga distancia de la marca Palomas de Paz es tal que la variable R que mide la distancia entre donde cayó un misil y su destino original (en km.), tiene la siguiente densidad sobre [0,1]:

$$f_R(r) = 2(1-r)$$

ER y Var(R). Calcula la probabilidad de que un misil caiga a menos de 100 metros de su destino original, lo que se considera como objetivo destruido. Si los vuelos de cada misil son independientes entre sí ¿ Cuántos misiles se tienen que lanzar en promedio para que el objetivo quede destruido?
Calculando ER:

$$\begin{aligned}
ER &= \int_{-\infty}^{\infty} r f_R(r) dr \\
&= \int_0^1 r f_R(r) dr \\
&= \int_0^1 2(r - r^2) dr \\
&= 2\int_0^1 (r - r^2) dr \\
&= 2\left(\frac{r^2}{2} - \frac{r^3}{3}\right)\Bigg|_0^1 \\
&= 2\left(\frac{1}{2} - \frac{1}{3}\right) \\
&= 2\left(\frac{1}{6}\right) \\
ER &= \frac{1}{3}
\end{aligned}$$

Calculando $Var(R)$, como ya se tiene cuanto vale ER, entonces se calculara $E(R^2)$.

$$\begin{aligned}
E(R^2) &= \int_{-\infty}^{\infty} r^2 f_R(r)dr \\
&= \int_0^1 r^2 f_R(r)dr \\
&= \int_0^1 2(r^2 - r^3)dr \\
&= 2\int_0^1 (r^2 - r^3)dr \\
&= 2\left(\frac{r^3}{3} - \frac{r^4}{4}\right)\Bigg|_0^1 \\
&= 2\left(\frac{1}{3} - \frac{1}{4}\right) \\
&= 2\left(\frac{1}{12}\right) \\
E(R^2) &= \frac{1}{6}
\end{aligned}$$

Entonces:

$$\begin{aligned}
Var(R) &= E(R^2) - (ER)^2 \\
&= \frac{1}{6} - \left(\frac{1}{3}\right)^2 \\
&= \frac{1}{6} - \frac{1}{9} \\
Var(R) &= \frac{1}{18}
\end{aligned}$$

Sea A un evento en el cual un misil impacte a menos de 100 metros de su destino original.

$$A = \{\omega \in R : R(\omega) \leq 0.1\}$$

entonces

$$\begin{aligned}
P(A) &= \int_0^{0.1} f_R(r)dr \\
&= \int_0^{0.1} 2(1-r)dr \\
&= 2\int_0^{0.1} (1-r)dr \\
&= 2\left(r - \frac{r^2}{2}\right)\Bigg|_0^{0.1} \\
&= 2\left(0.1 - \frac{(0.1)^2}{2}\right) \\
&= 2(0.095) \\
P(A) &= 0.19
\end{aligned}$$

Como el vuelo de cada misil es independiente, entoces R puede ser representado con una distribución geometrica ($R \sim Geo(p)$). Entonces, el valor esperado de la cantidad de misiles que te tienen que lanzar para que el objetivo quede destruido es:

$$\begin{aligned} E(A) &= \frac{1 - P(A)}{P(A)} \\ &= \frac{1 - 0.19}{0.19} \\ E(A) &= 4.2631 \end{aligned}$$

Problema 6

Se tiene una barra de longitud 1 metro y una posición p dada en la barra. Se rompe la barra en dos en un punto elegido al azar. Calcula la longitud promedio de la pieza de la barra que contiene p $(0 < p < 1)$.
Sea A todos los puntos posibles de elección.

$$A = \{p \in [0, 1]\}$$

Se tiene que para un punto p escogido de la barra, se obtienen las siguientes longitudes:

$$L(p) = \begin{cases} x & \text{para } 0 < p < x \\ 1 - x & \text{para } x < p < 1 \end{cases}$$

donde x es el punto donde se partio la barra. Entonces, el promedio de las logitudes qe contienen a p es:

$$\begin{aligned} E(L(p)) &= \int_A xL(p)dx \\ &= \int_0^p x(x)dx + \int_p^1 x(1-x)dx \\ &= \frac{x^2}{2}\Big|_0^p + \left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_p^1 \\ &= \frac{p^2}{2} + \frac{1}{2} - \frac{1}{3} - \frac{p^2}{2} + \frac{p^3}{3} \\ E(L(p)) &= \frac{p^3}{3} + \frac{1}{6} \end{aligned}$$

Problema 7

Define

$$f_{X,Y}(x, y) = \begin{cases} kyx & \text{con } 0 < x < 2, 0 < y < 1, 2y < x \\ 0 & \text{en otro caso} \end{cases}$$

Calcula el valor de k y P(3Y<X).
Calculando de k:

$$
\begin{aligned}
\int_0^1 \int_0^{2y} kxydxdy &= 1 \\
\int_0^1 ky \int_0^{2y} xdxdy &= 1 \\
\int_0^1 ky \left(\frac{x^2}{2}\right)\Big|_0^{2y} dy &= 1 \\
\frac{k}{2}\int_0^1 y(2y)^2 dy &= 1 \\
2k\int_0^1 y^3 dy &= 1 \\
2k\left(\frac{y^4}{4}\right)\Big|_0^1 &= 1 \\
\frac{k}{2} &= 1 \\
k &= 2
\end{aligned}
$$

Calculando $P(3Y < X)$, como $0 < x < 2$ y $0 < y < 1$, entonces se tiene que obtener el valor maximo para y, tal que $3Y < X$, esto sería

$$
\begin{aligned}
3Y &< X \\
3Y &< 2 \\
Y &< \frac{2}{3}
\end{aligned}
$$

entonces:

$$
\begin{aligned}
P(3Y < X) &= \int_0^{\frac{2}{3}} \int_0^{3y} 2xy\ dxdy \\
&= \int_0^{\frac{2}{3}} 2y \int_0^{3y} x\ dxdy \\
&= \int_0^{\frac{2}{3}} 2y \left(\frac{x^2}{2}\right)\Big|_0^{3y} dy \\
&= \int_0^{\frac{2}{3}} y(3y)^2 dy \\
&= \int_0^{\frac{2}{3}} 9y^3 dy \\
&= \frac{9}{4}y^4\Big|_0^{\frac{2}{3}} \\
&= \left(\frac{9}{4}\right)\left(\frac{2}{3}\right)^4 \\
&= \left(\frac{9}{4}\right)\left(\frac{16}{81}\right) \\
P(3Y < X) &= \frac{4}{9}
\end{aligned}
$$