Departamento de Ingeniería Civil Industrial
Probabilidades - IN3141
Profesor: Charles Thraves - Natalia Trigo
Aux's: Camila Carrasco - Monserrat Marchant - Leonardo Meneses - Sebastián Rojas

Auxiliar #4

Variable aleatoria discreta I

Resumen

- Variable Aleatoria Una función $X : \Omega \to \mathbb{R}$ es una **Variable Aleatoria**. Denotamos $R_X = \{x \in \mathbb{R} : X = x > 0\}$ como el rango de X: los valores que la variable aleatoria puede tomar.
- Función de masa de probabilidad Se define la **función de masa de probabilidad** (PMF: *Probability Mass Function*) para cada variable aleatoria X, como:
$$p_X(x) = P(X = x)\ \forall x \in R_X$$
- Función de distribución acumulada Se define la **función de distribución acumulada** (CDF: *Cumulative Distribution Function*) para cada variable aleatoria X, como:
$$F_X(x) = P(X \leq x)$$
- Propiedades de CDF Para toda variable aleatoria X, se tiene que:
  1. $F_X(x)$ es creciente en x.
  2. $$\lim_{x\to+\infty} F_X(x) = 1$$
  3. $$\lim_{x\to-\infty} F_X(x) = 0$$
  4. Para X variable aleatoria discreta, se tiene que para todo $x \in R_X$, existe un $\varepsilon > 0$ suficientemente pequeño para el cual
$$p_X(x) = F_X(x) - F_X(x - \varepsilon)$$
- Esperanza de una VA Para cada variable aleatoria, se define su **valor esperado o esperanza**, como:
$$\mathbb{E}[X] = \sum_{x\in R_X} x \cdot p_X(x)$$
  Observación:
  1. Puede ocurrir que $\mathbb{E}[X] = +\infty$.
  2. A veces se denota $\mathbb{E}[X] := \mu_X$
- Propiedades de la esperanza: Para cualquiera X, Y variables aleatorias, y $\lambda \in \mathbb{R}$ no aleatorio, tenemos que:
  1. Es lineal: $\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda\mathbb{E}[Y]$
  2. Si X, Y son independientes, entonces $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$
- Esperanza de una función de una v.a. Se tiene que:
$$E[g(X)] = \sum_{x\in R_X} g(x) \cdot p_X(x)$$
  Dato curioso: Esto es una propiedad y no una definición (para los curiosos, buscar *Law of the Unconscious Statician* [LOTUS].
- Varianza de una VA Se define, para una variable aleatoria, su **varianza**, como:
$$\mathrm{Var}(X) = E\left[(X - E[X])^2\right]$$
- Propiedades de la varianza Para cualquiera X, Y variables aleatorias, y $a, b \in \mathbb{R}$ no aleatorios, tenemos que:
  1. $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
  2. $Var(aX + b) = a^2Var(X)$
  3. Si X, Y son independientes, entonces $Var(X + Y) = Var(X) + Var(Y)$

  donde $\mathbb{E}[X^2]$ se calcula usando la esperanza de la función $g(t) = t^2$.
- Desviación estándar Se define, para una variable aleatoria, su **desviación estándar**, como:
$$SD(X) = \sqrt{\mathrm{Var}(X)}$$
- Coeficiente de variación Para una variable aleatoria X, definimos su **coeficiente de variación**, como
$$CV(X) = \frac{SD(X)}{\mathbb{E}[X]}$$

Distribuciones conocidas:

- **Bernoulli:** Sea $\mathrm{X} \in \{0,1\}, \mathrm{P}(\mathrm{X}=1)=\mathrm{p}, \mathrm{P}(\mathrm{X}=0)=1-\mathrm{p}, \mathrm{E}[\mathrm{X}]=\mathrm{p} \quad \mathrm{Var}(\mathrm{X})=\mathrm{p}(1-\mathrm{p})$
- **Binomial:** Sea $\mathrm{X} \sim \mathrm{bin}(n,p), \mathrm{P}(\mathrm{X}=\mathrm{k})=\left(\begin{array}{c} n \\ k \end{array}\right) p^k(1-p)^{n-k}, \mathrm{E}[\mathrm{X}]=\mathrm{np}, \mathrm{Var}[\mathrm{X}]=\mathrm{np}(1-\mathrm{p})$
- **Hipergeométrica:** $\mathrm{X} \sim$ Hipergeométrica $(\mathrm{N},\mathrm{K},\mathrm{n})$ con $\mathrm{N},\mathrm{K},\mathrm{n} \in \mathbb{Z}_+, \mathrm{K},\mathrm{n} \leq \mathrm{N}, R_x = \{\max\{0, n+K-N\}, \ldots, \min\{K,N\}\}$, con PMF

$$P_X(k)=\frac{\left(\begin{array}{c} K \\ k \end{array}\right)\left(\begin{array}{c} N-k \\ n-k \end{array}\right)}{\left(\begin{array}{c} N \\ n \end{array}\right)}$$

$$\text{con}\,\mathrm{E}[\mathrm{X}]=\mathrm{n}\frac{K}{N}\text{y}\,\mathrm{Var}[\mathrm{X}]=\mathrm{n}\frac{K}{N}\frac{(N-K)}{N}\frac{N-n}{N-1}$$

Pregunta 1

Se lanza un dado equilibrado con n caras enumeradas de 1 a n, y se anota el resultado obtenido. El procedimiento se repite hasta que se obtiene un resultado que ya se anotó en algún lanzamiento previo. Sea X la variable aleatoria que denota la cantidad total de lanzamientos.

a) ¿Cuál es el rango de X?

Solución

Viendo el caso más corto, este es aquel donde se lanza el dado por primera vez (notar que es imposible que se repita el resultado en el primer lanzamiento, pues no hay lanzamientos previos con los cual comparar) y luego en el segundo lanzamiento aparece la misma cara que en el primero. Por lo tanto son 2 lanzamientos.

El caso más largo corresponde a aquel donde se lanza el dado n veces y nunca se repitió alguna cara, pero en el lanzamiento $n+1$ es 100% seguro que se repetirá la cara, pues no quedan caras nuevas por salir. Por lo tanto se ejecutan $n+1$ lanzamientos.

Entonces $R_X = \{2, 3, ..., n+1\}$. Notar que este rango es discreto pues no pueden realizarse fracciones de lanzamientos (ej: 2.5 lanzamientos es imposible)

b) Calcula su función de probabilidad p_X si $n = 2$ (puede interpretar que es una moneda en vez de un dado).

Solución

Si $n = 2$, entonces $R_X = \{2, 3\}$.

$$\begin{aligned} \mathbb{P}(X=2) &= \frac{1}{2} \\ \mathbb{P}(X=3) &= \frac{1}{2} \end{aligned}$$

Cuando $X = 2$ la probabilidad es 0.5, viéndolo con un ejemplo, si en el primer lanzamiento salió sello, la probabilidad que vuelva a salir sello en el segundo (y por lo tanto de que se repita) es de 0.5.

Siguiendo con el mismo ejemplo, en $X = 3$ la probabilidad es 0.5 pues la probabilidad de que en el segundo lanzamiento **no** salga sello es 0.5, luego en el tercer lanzamiento si o si se repetirá alguna de las caras que salieron previamente y termina el juego.

Notar que como en toda PMF, esta debe sumar 1.

c) Calcule su función de probabilidad p_X para el caso genérico.

Solución

Veamoslo con ejemplos hasta encontrar una secuencia, si $X = 2$, eso significa que salió repetida la misma cara que en el primer lanzamiento, como cada lanzamiento es independiente, la probabilidad que esa cara haya sido la misma es:

$$\mathbb{P}(X = 2) = \frac{1}{n}$$

En el caso de $X = 3$, esto significa que en el segundo lanzamiento no se repitió la cara del primero, es decir, salió una de las $n - 1$ caras diferentes. En el tercer lanzamiento salió una cara repetida, pero ahora se tiene dos posibilidades de cara repetida, esto es igual a que:

$$\mathbb{P}(X = 3) = \frac{n-1}{n} \cdot \frac{2}{n}$$

En $X = 4$, el segundo y tercer lanzamiento no salió una cara repetida, es decir, en el segundo salió una de $n - 1$ caras diferentes y en el tercero salio una de las $n - 2$ caras diferentes. En el cuarto lanzamiento se repitió una cara, es decir, salió una de las 3 caras diferentes que habían salido antes.

$$\begin{aligned} \mathbb{P}(X = 4) &= \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \frac{3}{n} \\ &= \frac{(n-1)!}{(n-3)!} \cdot \frac{3}{n^3} \end{aligned}$$

Generalizando para $X = x$, con $x \in R_X$, la pmf queda expresada como:

$$\mathbb{P}(X = x) = \frac{(n-1)!}{(n-(x-1))!} \cdot \frac{(x-1)}{n^{x-1}}$$

Donde el primer termino son los elementos que no se repiten y el segundo termino cuando se repite una cara y se termina el juego.

Pregunta 2

La empresa *Júpiter* produce termómetros infrarrojos. Cada termómetro producido queda bien calibrado con una probabilidad de p, en cuyo caso el termómetro, al ser usado, siempre indicará la temperatura correcta (o real) del objeto al cual se le mide su temperatura. En caso de que el termómetro quede mal calibrado, éste entregará un valor desviado por δ grados Celsius respecto a la temperatura real del objeto. Asuma que en ambos casos, donde la lectura de la temperatura se desvía por **sobre** o **bajo** δ grados Celsius, son equiprobables (en el caso de que el termómetro quede mal calibrado[1]). Llamaremos al parámetro δ el *error*.

La empresa *Bedebueno* testea los termómetros producidos por la empresa *Júpiter*. Específicamente, mide la temperatura de un objeto que está a una temperatura (real) de a grados Celsius con un termómetro producido por la empresa *Júpiter*. Llame X a la temperatura que entrega la lectura con un termómetro producido por la empresa *Júpiter*.

Asuma que $a > 0$, $\delta > 0$, y $p \in (0, 1)$.

(a) Entregue la CDF de la VA X. Además, grafíquela.

Solución:

$$F_X(x) = \begin{cases} 0 & \text{si } x < a - \delta \\ \frac{1-p}{2} & \text{si } a - \delta \leq x < a \\ \frac{1+p}{2} & \text{si } a \leq x < a + \delta \\ 1 & \text{si } a + \delta \leq x \end{cases} \tag{1}$$

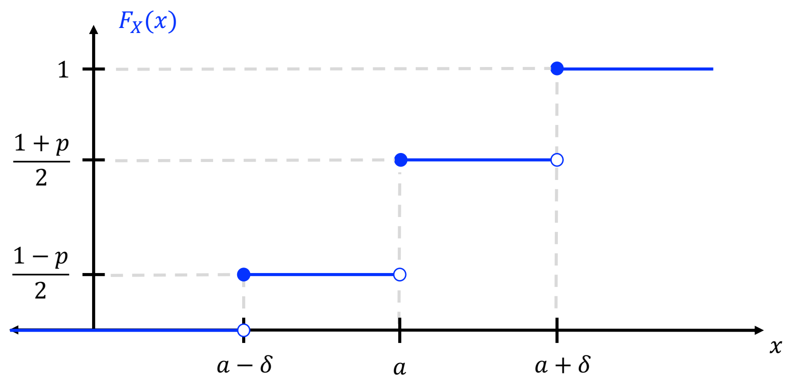

Figure 1: CDF

(b) *Bedebueno* está implementando una certificación de calidad, la cual se entrega a la empresa que produce los termómetros, en el caso de que esta última cumpla con el criterio establecido. En particular, dicho criterio consiste en que los termómetros producidos por la empresa, al medir la temperatura de un objeto a temperatura real de a grados Celsius, deben tener una desviación estándar menor o igual a $\delta/2$. Indique las condiciones de los parámetros del problema (p, a, y/o δ) de tal modo que *Júpiter* reciba la certificación.

Solución:

Primero necesitaremos la PMF. Notamos que:

$$p_X(x) = \begin{cases} \frac{1-p}{2} & \text{si } x = a - \delta \\ p & \text{si } x = a \\ \frac{1-p}{2} & \text{si } x = a + \delta \end{cases} \tag{2}$$

Tenemos que calcular ahora la desviación estándar, para eso, computamos primero la varianza, la cual la podemos calcular de dos formas:

- **Forma 1:** $\mathrm{Var}(X) = E[X^2] - E[X]^2$.

$$\begin{aligned} E[X] &= \frac{1-p}{2}(a-\delta) + pa + \frac{1-p}{2}(a+\delta) && (3) \\ &= a && (4) \\ E[X^2] &= \frac{1-p}{2}(a-\delta)^2 + pa^2 + \frac{1-p}{2}(a+\delta)^2 && (5) \\ &= (1-p)(a^2+\delta^2) + pa^2 && (6) \end{aligned}$$

Luego

$$\begin{aligned} Var[X] &= E[X^2] - E[X]^2 && (7) \\ &= (1-p)(a^2+\delta^2) + pa^2 - a^2 && (8) \\ &= (1-p)\delta^2 && (9) \end{aligned}$$

- **Forma 2:** $\mathrm{Var}(X) = E[(X - E[X])^2]$.

$$\begin{aligned} E[X] &= \frac{1-p}{2}(a-\delta) + pa + \frac{1-p}{2}(a+\delta) && (10) \\ &= a && (11) \\ \mathrm{Var}(X) = E[(X-E[X])^2] &= \frac{1-p}{2}(a-\delta-a)^2 + p(a-a)^2 + \frac{1-p}{2}(a+\delta-a)^2 && (12) \\ &= (1-p)\delta^2 && (13) \end{aligned}$$

Luego, $\sigma_X = \delta\sqrt{1-p}$. Entonces

$$\begin{aligned} \sigma_X \leq \frac{\delta}{2} &\Leftrightarrow \delta\sqrt{1-p} \leq \frac{\delta}{2} && (14) \\ &\Leftrightarrow \sqrt{1-p} \leq \frac{1}{2} && (15) \\ &\Leftrightarrow 1-p \leq \frac{1}{4} && (16) \\ &\Leftrightarrow \frac{3}{4} \leq p && (17) \end{aligned}$$

(c) Suponga que la empresa *Júpiter* está considerando arrendar una tecnología que le permita controlar el *error* de la lectura del termómetro. El precio de arriendo de dicha tecnología es de $\$\lambda/\delta$ por cada termómetro producido con un *error* de δ (donde $\lambda > 0$ es un parámetro fijo y conocido). Además, considere que ahora *Júpiter* debe pagar en reparación un monto de $\$c \cdot \delta$ por cada termómetro malo que tenga una lectura de δ grados Celsius por **sobre** o **bajo** la temperatura real del objeto al cual se le mide su temperatura.

(i) Dado un parámetro fijo de *error* igual a $\delta > 0$, obtenga el costo esperado de producir un termómetro (incluya todos los costos asociados, es decir, de arriendo de tecnología y de reparación).
Hint: Le podría ser de utilidad escribir el costo como una expresión que dependa de la VA X.

Solución:
El costo esperado, dado un valor fijo del error δ, Y lo podemos escribir como $Y = g(X) = \frac{\lambda}{\delta} + c|X-a|$. Luego

$$\begin{aligned} E[Y] &= E\left[\frac{\lambda}{\delta} + c|X-a|\right] \\ &= \frac{\lambda}{\delta} + cE[|X-a|] \\ &= \frac{\lambda}{\delta} + c\left(\frac{1-p}{2}|a-\delta-a| + p|a-a| + \frac{1-p}{2}|a+\delta-a|\right) \\ &= \frac{\lambda}{\delta} + c(1-p)\delta \end{aligned} \tag{18}$$

(ii) Considerando que *Júpiter* puede escoger cualquier valor no negativo para el parámetro de *error* a contratar con la nueva tecnología, indique el valor óptimo que debería utilizar.

Solución:
Simplemente debemos minimizar la Ecuación (18) con respecto a δ. Derivando e igualando a cero: $-\frac{\lambda}{\delta^2} + c(1-p) = 0$, luego $\delta^* = \sqrt{\frac{\lambda}{c(1-p)}}$. Notamos que es un mínimo, pues la segunda derivada es $2\lambda/\delta^3 > 0$.

Pregunta 3

Un jugador de emboque se inscribe a un torneo durante las fiestas patrias. El jugador tiene 10 intentos, y la probabilidad de acertar es de 0.2 (cada intento es independiente).

a) Calcule la PMF de X
Solución:
Se tienen 10 intentos, en donde la probabilidad de aciertos es de p = 0.2, independientes entre sí. Se define el evento X como el número de aciertos, por ende se sabe que:
X es una binomial con $n = 10, p = 0.2$ $X \sim \text{Bin}(10, 0.2)$

$$\rightarrow P_x(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{10}{k} 0.2^k 0.8^{10-k}$$
$$\text{Para todo } \; k \in R_x = \{0, \cdots, n\} = \{0, ..., 10\}$$

b) ¿Cuál es la probabilidad de no acertar en ningún intento?
Solución:
No acertar en ningún intento nos dice que $X = 0$, reemplazando en la PMF de X cuando $X = 0$ nos entrega:

$$\begin{aligned} P(X=0) &= \binom{10}{0} 0.2^0 0.8^{10} = \frac{10!}{10!0!} \cdot 1 \cdot 0.8^{10} \\ &= 0.107 \end{aligned}$$

c) ¿Cuál es la probabilidad de que acierte más veces de las que falla?
Solución:

$$P(X \geq 6) = \sum_{k=6}^{10} P(X = k) = 0.0064$$

d) Calcule la Esperanza y Varianza de X.

Solución:

Dado que la Variable Aleatoria X distribuye binomial utilizamos la definición de varianza y esperanza asociada, de esta forma se obtiene lo siguiente:

$$E[X] = n \cdot p = 10 \cdot 0.2 = 2$$
$$\text{Var}[X] = n \cdot p(1-p) = 10 \cdot 0.2 \cdot 0.8 = 1.6$$

e) **Propuesto:** Suponga que el jugador tiene que pagar \$3 para entrar al torneo y que gana \$2 por cada acierto. Sea Y su ganancia, encuentre su Esperanza y Varianza.

Solución:

Ahora el jugador paga \$3 para entrar al torneo y gana \$2 por cada acierto, por lo que la función de ganancia, Y, se encuentra definida en función de la variable aleatoria X.

Función de ganancia:

$$Y = 2 \cdot X - 3$$

Definición de esperanza de una función:

$$\begin{aligned} E[Y] = E[g(X)] = E[2X-3] &= E[2X] - E[3] \\ &= 2E[X] - 3 = 2 \cdot 2 - 3 = 1 \end{aligned}$$

Varianza de la ganancia del jugador:

$$\text{Var}[Y] = \text{Var}[2X-3] = \text{Var}[2 \cdot X] = 2^2 \,\text{Var}[X]$$

$$= Var[Y] = 4 \cdot 1.6 = 6.4$$

f) **Propuesto:** Ahora suponga que no tiene que pagar para entrar al torneo y que su ganancia es el número de aciertos al cuadrado. Si Z es su ganancia, encuentre la Esperanza de Z.

Solución:

Modificando la función de ganancia, ahora dada por el número de aciertos al cuadrado, obtenemos una nueva esperanza de Z, con Z la ganancia del jugador.

$$\begin{aligned} Z = X^2 \rightarrow E[Z] = E\left[X^2\right] = \text{Var}(X) + E(X)^2 &= 1.6 + 4 \\ &= 5.6 \end{aligned}$$