

Trabajo práctico: unidad 4

Probabilidad y estadística

Martín Rossi

1.

a.

$$\begin{aligned} F(4) &= P(X \leq 4) \\ &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= 0.32 + 0.35 + 0.18 + 0.08 + a \\ &= 0.93 + a \end{aligned}$$

$$0.97 = F(4) = 0.93 + a \implies a = 0.04$$

Como $S = \{0, 1, 2, 3, 4, 5, 6\}$, $F(6) = 1$

$$1 = F(6) = F(4) + P(X = 5) + P(X = 6) = 0.97 + b + 0.01 = b + 0.98 \implies b = 0.02$$

b.

i. $P(X \leq 4) = F(4) = 0.97$

ii. $P(X \geq 5) = 1 - F(4) = 0.03$

c.

$$\begin{aligned} E(X) &= \sum_{i=0}^6 i * P(X = i) \\ &= 0 * 0.32 + 1 * 0.35 + 2 * 0.18 + 3 * 0.08 + 4 * 0.04 + 5 * 0.02 + 6 * 0.01 \\ &= 1.27 \end{aligned}$$

$$\begin{aligned} E(X^2) &= \sum_{i=0}^6 i^2 * P(X = i) \\ &= 0 * 0.32 + 1 * 0.35 + 4 * 0.18 + 9 * 0.08 + 16 * 0.04 + 25 * 0.02 + 36 * 0.01 \\ &= 3.29 \end{aligned}$$

$$SD(X) = \sqrt{Var(X)} = \sqrt{E(X^2) - E(X)^2} = \sqrt{3.29 - 1.27^2} = \sqrt{3.29 - 1.6129} = 1.2950$$

En promedio habrá 1.27 interrupciones diarias. Por el desvío estándar se puede decir que el 68.2% de los días habrá entre 0 y 2.565 interrupciones (1.27 ± 1.2950), el 95.4% habrá entre 0 y 2.9470 interrupciones ($1.27 \pm 2 * 1.2950$), siendo 0 el mínimo posible.

d. Son parámetros, porque la variable aleatoria está tomada de toda la población, que serían la cantidad de interrupciones diarias todos los días que la compañía estuvo activa.

2.

Sin técnica de corrección de errores: ϵ , como dice el enunciado.

Con corrección de errores: Se define la variable aleatoria

X : cantidad de bits mal interpretados

$X \sim \text{Bin}(3, \epsilon)$

Habría un error si $X \geq 2$

$$P(X \geq 2) = P(X = 2) + P(X = 3) = \binom{3}{2} * \epsilon^2 * (1 - \epsilon) + \binom{3}{3} * \epsilon^3 = \epsilon^2 * (3 + 2 * \epsilon)$$

Si se toma $\epsilon = 0.1$, con la técnica la probabilidad de cometer error sería

$0.1^2 * (3 + 2 * 0.1) = 0.032$, aproximadamente un tercio de error que sin usar la técnica.

Pero tiempo de transmisión sería el triple porque habría que enviar cada bit repetido 3 veces.

Si $\epsilon = 0.05$, usar la técnica reduciría la probabilidad de error a

$0.05^2 * (3 + 2 * 0.05) = 0.0077$, un 15% de la probabilidad sin la técnica. Y el tiempo sería también el triple.

Mientras menor sea la probabilidad de error, más conveniente será utilizar la técnica de corrección, pero menos necesaria será porque ya la probabilidad es baja.

3. Como se considera que la cantidad de aspirantes es muy grande, el porcentaje de avanzados no cambiará a medida que se extraigan para hacer las entrevistas.

a.

X : número de entrevistas necesarias hasta encontrar el primer aspirante avanzado

$X \sim \text{Geom}(p), p = 0.3$

$$P(X = 5) = q^4 * p = 0.7^4 * 0.3 = 0.0720$$

b.

Y : número de entrevistas necesarias hasta encontrar el quinto aspirante avanzado

$Y \sim \text{Pascal}(5, p), p = 0.3$

$$P(Y = 10) = \binom{9}{4} * 0.3^5 * 0.7^5 = 0.05146$$