Ana Llenado - 14/09/23

1) a) Qué es una variable aleatoria? Cuando decimos que una v.a. es discreta o continua?

- Una variable aleatoria es una función que asigna un valor numérico a cada resultado posible de un experimento aleatorio.
- Es cualquier subconjunto del espacio muestral S.

V.a. Discreta

Una v.a. discreta toma un conjunto finito o infinito numerable de valores posibles.

V.a. Continua

Puede tomar cualquier valor dentro de un intervalo de números $\mathbb{R}$.

Ej: tiempo, altura, peso, ingresos.

b) Define Función de probabilidad de masa, enunciando las propiedades que debe satisfacer

- Sea X una v.a. discreta. La función de probabilidad de masa de X es una función $p: I \subseteq \mathbb{R} \rightarrow [0,1]$ que satisface:

1. $P(x_i) = P(X = x_i)$
2. $P(x_i) \geq 0 \ \forall i$
3. $\sum_{i=1}^{n} P(x_i) = 1$

c) Define Función de Distribución Acumulada, relacionándola con la función anterior y con la probabilidad. Ver Teorema 4.1

La Función de Distribución Acumulada de una v.a. discreta X, denotada por F(x), está dada por:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} P(x_i)$$

donde P(x_i) es la función de probabilidad de masa de X.

$F(x)$ se puede relacionar con el calculo de la probabilidad de que X se encuentre dentro de un intervalo $(a,b]$ con $a<b$ dado que:

$$P(a < x \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

d) Cómo calcula la esperanza de una v.a. discreta?

Esperanza de una v.a. discreta

Sea X una v.a. discreta y $p(x)$ su función de probabilidad de masa

$$E(x) = \sum x . p(x)$$

Esperanza de una v.a. continua

Sea X una v.a. continua y $f(x)$ su funcion de densidad.

$$E(x) = \int_{-\infty}^{\infty} x . f(x) \, dx$$

e) Dé un ejemplo de v.a. discreta. Cuál es su funcion de distribucion acumulada?

Un ejemplo de v.a. discreta es el lanzamiento de un dado numerado (equilibrado). X representa el numero de la cara que queda hacia arriba despues de lanzarlo.

$S = \{1, 2, 3, 4, 5, 6\}$. $F(x) = P(X \leq x)$

$$F(x) \begin{cases} 1/6 & x \leq 1 & P(X \leq 1) \\ 2/6 & x \leq 2 & P(X \leq 2) \\ 3/6 & x \leq 3 & P(X \leq 3) \\ 4/6 & x \leq 4 \\ 5/6 & x \leq 5 \\ 1 & x \leq 6 \end{cases}$$

2) Cuál modelo es el que mejor describiría a la v.a. X en cada una de las siguientes situaciones? Justifique su respuesta especificando en cada caso el o los parametros del modelo elegido.

a) De una estacion parte un tren cada 30 minutos. Un viajero llega de improviso. Sea X = tiempo de espera hasta la proxima partida

- En este caso, la v.a. X representa el tiempo de espera hasta la proxima partida de un tren. Dado que los trenes parten cada 30 minutos, el tiempo de espera es continuo, un modelo adecuado seria el exponencial.

Parámetro: La tasa de llegada (λ) que es inversa al tiempo promedio entre llegadas.

En este caso $\lambda = \frac{1}{30} = \frac{1}{\mu}$ $\mu = 30$ min
$\downarrow$
tiempo promedio de espera.

La distribucion exponencial se utiliza comunmente para modelar el tiempo entre eventos en un proceso de Poisson, donde los eventos ocurren de manera continua e independiente a una tasa constante.

b) En cierto servicio de emergencias se solicita en promedio 8 ambulancias por día. Sea X = numero de ambulancias solicitadas en las proximas 12 hs.

- En este escenario la v.a. X representa el numero de ambulancias solicitadas en un tiempo fijo. Este caso se ajusta a una distribucion de Poisson.

Parametro: La tasa de ocurrencia (λ), que es el promedio de eventos por unidad de tiempo.

En este caso $\lambda = 8$ por día, entonces para 12hs seria $\lambda = \frac{8}{2} = 4$

La funcion de probabilidad de distribución:

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$$

ⓒ Un producto electrónico contiene 40 circuitos integrados, que funcionan independientemente uno de otro. La probabilidad de que alguno de ellos falle es de 0,01 Sea X = numero de circuitos integrados defectuosos.

- Dado que hay un numero fijo de circuitos y cada uno funciona de manera independiente, y siendo que hay solo dos resultados posibles Falla o No falla, se puede modelar este escenario utilizando una distribucion binomial.

Parametro

La distribucion binomial tiene 2 parametros: n (numero de ensayos) y p (probabilidad de exito)

$\Rightarrow n = 40 \wedge p = 0{,}01$

Funcion de probabilidad de masa

$$p(x) = P(X = x) = \binom{n}{x} p^{x} . (1-p)^{n-x}$$

3) Enuncie alguna de las versiones del Teorema Central del Limite y mencione algún tema de esta asignatura en el que hayamos requerido hacer uso de este resultado.

- Un tema en el que se requiere el uso del TCL es en la inferencia estadística, especialmente en la estimación de intervalos de confianza, y en las pruebas de hipótesis

Por ejemplo, al estimar un intervalo de confianza para la media de una población o al realizar pruebas de hipótesis sobre la media poblacional, se utiliza el TCL para justificar el uso de la distribución normal para aproximar la distribución muestral de la media.

4) a) En que consiste la estimación puntual de un parámetro poblacional?

- Consiste en utilizar un solo valor, llamado estimador puntual, para hacer una conjetura o estimación acerca del valor del parámetro en la población. Es decir, se busca proporcionar una única estimación numérica que sirva como mejor conjetura del valor real del parámetro de interés.

b) Mencione un método de estimación puntual y desarróllelo

- Método de Máxima Verosimilitud:

Este método busca encontrar el valor del parámetro que maximiza la función de verosimilitud, que es la probabilidad de observar los datos, dados ciertos valores del parámetro.

① Escribir la función de verosimilitud

② Tomar el Logaritmo de la función de Verosimilitud

③ Derivar con respecto al parámetro e igualar a cero.

④ Verificar que es un Máximo utilizando la segunda derivada o mediante criterios de concavidad.

⑤ El valor que maximiza la función de verosimilitud se convierte en el estimador puntual.

c) Qué significa que un estimador puntual sea insesgado?

- Un estimador puntual se considera insesgado si su valor esperado o media muestral es igual al valor del parámetro que se está esperando.

Formalmente, para un estimador $\hat{\theta}$ de un parámetro θ:

$$E\,\hat{\theta} = \theta$$

Un estimador insesgado, en promedio, a lo largo de muchas muestras, no sobreestima ni subestima el parámetro poblacional.

d) Si tuviera dos estimadores insesgados de un parámetro, cuál elegiría?

Elegimos el estimador insesgado de mínima varianza.

5) En una prueba de hipótesis (de una cola) para la media poblacional y un nivel de significancia de α:

a) Cuándo utilizaría un estadístico test t y cuándo uno z?

La elección entre uno y otro depende de la disponibilidad de información sobre la desviación estándar poblacional (σ)

① Estadístico de Test t:

- Se utiliza cuando la desviación estándar poblacional (σ) no es conocida y se estima a partir de la muestra (s).
- adecuado para muestras pequeñas
- La distribución de muestreo se aproxima a una distribución t de student.

② Estadístico de Test z:

- Se utiliza cuando la desviación estándar poblacional (σ) es conocida.
- Adecuado para muestras grandes ($n \geq 30$).
- La distribución de muestreo se aproxima a una normal estándar.

3) b) Plantee un problema donde haya que realizar una prueba de hipotesis de 1 cola, especificando hipótesis nula, alternativa, nivel de significancia. Describa detalladamente cuales son los pasos.

- Supongamos que una empresa afirma que el tiempo promedio de entrega de sus productos es de hasta 4 dias. Sin embargo un gerente sospecha que el tiempo promedio de entrega es mayor.

① Establecer las hipotesis

$H_0 : \mu \leq 4$

$H_1 : \mu > 4$

② seleccionar el nivel de significancia (α).

Por ejemplo $\alpha = 0.05$

③ seleccionar el estadistico de prueba

Dado que estamos comparando la media de una muestra con un valor especifico, y no conocemos la desviacion estandar poblacional, se usaria el estadistico t.

$$\Rightarrow t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$\bar{X}$ = media muestral

μ_0 = valor bajo hipotesis nula

S = desviacion estandar muestral

n = tamaño de la muestra

④ Determinar la region critica

En este caso al ser una prueba unidireccional (mayor que), la region critica estaria en el extremo derecho de la distribucion t con n-1 grados de libertad.

⑤ Calcular el estadistico de prueba y compararlo con el valor critico si el estadistico de prueba cae en la region critica, rechazaremos la hipotesis nula.