Probabilidades y Estadística

Clase 6
Variable Aleatoria II

Nicolás Araya Caro

Universidad Diego Portales
Escuela de Informática y Telecomunicaciones

14 de septiembre de 2023

Contenidos

Contenidos

Distribución Bernoulli

La distribución Bernoulli es una distribución de probabilidad **Dicotómica** o **Binaria**, donde toma valor 1 para éxitos con probabilidad p y 0 para fracasos con probabilidad 1-p.
Su función de probabilidad es:

$$f(x) = p(1-p)^{1-x}$$

Escrito de otra forma:

$$f(x;p) = \begin{cases} p \text{ si}, & p = 1 \\ (1-p) \text{ si}, & p = 0 \\ 0 \text{ si}, & \text{En otro caso} \end{cases}$$

La Esperanza $E[X] = p = \mu$ y la Varianza $Var[X] = p(1-p) = pq$, con q igual a 1-p

Contenidos

Condiciones experimentales

Existen muchos experimentos que se ajustan exacta o aproximadamente a la siguiente lista de requerimientos:

1. El experimento consta de una secuencia de n experimentos más pequeños llamados ensayos, donde n se fija antes del experimento.
2. Cada ensayo puede dar por resultado uno de los mismos dos resultados posibles (ensayos dicotómicos), los cuales se denotan como éxito (E) y falla (F).
3. Los ensayos son independientes, de modo que el resultado en cualquier ensayo particular no influye en el resultado de cualquier otro ensayo.
4. La probabilidad de éxito es constante de un ensayo a otro; esta probabilidad se denota por p.

Definición:

Cumpliendo las 4 condiciones anteriores, es un experimento Binomial.

Como la función masa de probabilidad de una variable aleatoria binomial X depende de los dos parámetros n y p, la función masa de probabilidad se denota por b(x); n, p):

$$b(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{si,} \quad x = 0, 1, 2 \ldots n \\ 0 \text{ si,} & \text{En otro caso} \end{cases}$$

La Esperanza $E[X] = np$ y la Varianza $Var[X] = np(1-p)$.

Ejemplo

Se cree que 10 % de los paquetes entrantes que llegan a cierto servidor se **pierden**. Se seleccionan 10 paquetes entrantes al azar. ¿Cual es la probabilidad de que 4 de estos paquetes se pierdan?.

Respuesta: calculamos cuando de 10 paquetes, se pierden 4

$$f(4) = \binom{10}{4}(0.9)^4(1-0.9)^{10-4} = \frac{10!}{4!(10-4)!}(0.9)^4(0.1)^6$$

Distribución Acumulada

Para X $\sim$ Bin(n,p), la función de distribución acumulativa será denotada por:

$$P(X \leq x) = B(x; n, p) = \sum_{y=0}^{x} b(y; n, p) \quad \text{con x= 1,2,3...n}$$

La Esperanza $E[X] = np$ y la Varianza $Var[X] = np(1-p)$.

Ejemplo

Suponga que 20 % de todos los ejemplares de un libro de texto particular no pasan una prueba de resistencia de encuadernación. Sea X el número entre 15 ejemplares seleccionados al azar que no pasan la prueba. Entonces X tiene una distribución binomial con n = 15 y p = 0.2.

1. La probabilidad de que cuando mucho 8 no pasen la prueba es:

$$P(X \leq 8) = \sum_{y=0}^{8} b(y; 15, 0.2) = 0.999$$

2. La probabilidad de que exactamente 8 fallen es:

$$P(X = 8) = P(X \leq 8) - P(X \leq 7) = b(8; 15, 0.2) - b(7; 15, 0.2) = 0.003$$

Continuación

3 La probabilidad de que por lo menos 8 fallen:

$$P(X \geq 8) = 1 - P(X \leq 7) = 1 - B(7; 15, 0.2) = 1 - 0.996 = 0.004$$

4 Finalmente, la probabilidad de que entre 4 y 7, inclusive, fallen es:

$$\begin{aligned} P(4 \leq X \leq 7) &= P(X = 4, 5, 6 \text{ o } 7) = P(X \leq 7) - P(X \leq 3) \\ &= B(7; 15, 0.2) - B(3; 15, 0.2) \\ &= 0.996 - 0.648 = 0.348 \end{aligned}$$

Contenidos

Contenidos

Condiciones experimentales

Las suposiciones que conducen a la distribución hipergeométrica son las siguientes:

- La población o conjunto que se va a muestrear se compone de N individuos, objetos o elementos (una población *finita*).
- Cada individuo puede ser caracterizado como éxito (E) o falla (F) y hay M éxitos en la población.
- Se selecciona una muestra de n individuos sin reemplazo de tal modo que cada subconjunto de tamaño n es igualmente probable de ser seleccionado.

La variable aleatoria de interés es $X =$ el número de éxitos en la muestra. La distribución de probabilidad de X depende de los parámetros n, M y N, así que se desea obtener $P(X = x) = h(x; n, M, N)$.

Ejemplo

En una fábrica, existen lotes de 40 componentes donde cada uno se denominan aceptables si no contienen más de tres componentes defectuosos. El procedimiento para muestrear el lote es la selección de cinco componentes al azar y rechazar el lote si se encuentra un componente defectuoso. ¿Cuál es la probabilidad de que se encuentre exactamente un componente defectuoso en la muestra si hay tres componentes defectuosos en todo el lote?

Utilizando la distribución hipergeométrica con n=5, N=40, M=3 y x=1, entonces:

$$h(1; 40, 5, 3) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3$$

Distribución Hipergeométrica

Si X es el número de éxitos (E) en una muestra completamente aleatoria de tamaño *n* extraída de la población compuesta de M *éxitos* y (N - M) *fallas*, entonces la distribución de probabilidad de X llamada **distribución hipergeométrica**, es:

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

con x un entero que satisface máx(0, n - N + M) $\leq$ x $\leq$ mín(n, M).

La media y la varianza de la variable aleatoria hipergeométrica X cuya función masa de probabilidad es h(x; n, M, N) son:

$$E(X) = n \cdot \frac{M}{N}$$

$$V(X) = n \cdot (\frac{N-n}{N-1}) \cdot \frac{M}{N} \cdot (1 - \frac{M}{N})$$

Contenidos

Distribución de Poisson

Se dice que una variable aleatoria X tiene una distribución de Poisson con parámetro $\lambda (\lambda > 0)$ si la función masa de probabilidad de X es:

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad x = 0, 1, 2...$$

Ejemplo

Sea X el número de criaturas de un tipo particular capturadas en una trampa durante un periodo determinado. Suponga que X tiene una distribución de Poisson con $\lambda = 4.5$, así que en promedio las trampas contendrán 4.5 criaturas. La probabilidad de que una trampa contenga exactamente cinco criaturas es:

$$p(X = 5) = \frac{e^{-4.5}(4.5)^5}{5!} = 0.1708$$

La probabilidad de que una trampa contenga cuando mucho cinco criaturas:

$$P(X \leq 5) = \sum_{x=0}^{5} = e^{-4.5}\left[1 + 4.5 + \frac{(4.5)^2}{2!} + \ldots + \frac{(4.5)^5}{5!}\right] = 0.1708$$

Distribución de Poisson como límite

Proposición

Suponga que en la función masa de probabilidad binomial b(x; n, p), si $n \to \infty$ y $p \to 0$ de tal modo que np tienda a un valor $\lambda > 0$. Entonces b(x; n, p) $\to p(x; \lambda)$

De acuerdo con esta proposición, en cualquier experimento binomial en el cual n es grande y p es pequeña, $b(x; n, p) \approx p(x; \lambda)$, donde $\lambda = np$ Como regla empírica, esta aproximación puede ser aplicada con seguridad si n $>$ 50 y np $<$ 5.

Nota

Si X tiene una distribución de Poisson con parámetro λ, entonces E(X) $=$ V(X) $=\lambda$.

Una aplicación muy importante de la distribución de Poisson surge en conexión con la ocurrencia de eventos de algún tipo en el transcurso del tiempo. Se hace la siguiente suposición sobre la forma en que los eventos de interés ocurren:

1. Existe un parámetro $\alpha > 0$ de tal modo que durante cualquier intervalo de tiempo corto Δt, la probabilidad de que ocurra exactamente un evento es $\alpha \cdot \Delta t + o(\Delta t)*$
2. La probabilidad de que ocurra más de un evento durante Δt es $o(\Delta t)$ [la que junto con la suposición 1, implica que la probabilidad de cero eventos durante Δt es $1 - \alpha \cdot \Delta t - o(\Delta t)$].
3. El número de eventos ocurridos durante este intervalo de tiempo Δt es independiente del número ocurrido antes de este intervalo de tiempo.

Proposición

Informalmente, la suposición 1 dice que durante un corto intervalo de tiempo, la probabilidad de que ocurra un solo evento es aproximadamente proporcional a la duración del intervalo de tiempo, donde α es la constante de proporcionalidad. Ahora sea $P_k(t)$ la probabilidad de que k eventos serán observados durante cualquier intervalo de tiempo particular de duración t.

Proposición

$P_k(t) = e^{-\alpha\lambda} \cdot (\alpha t)^k / k!$, de modo que el número de eventos durante un intervalo de tiempo de duración t es una variable de Poisson con parámetro $\lambda = \alpha t$. El número esperado de eventos durante cualquier intervalo de tiempo es entonces αt así que el número esperado durante un intervalo de tiempo unitario es α

Ejemplo

Suponga que llegan pulsos a un contador a un ritmo promedio de seis por minuto, así que $\alpha = 6$ Para determinar la probabilidad de que en un intervalo de 0.5 min se reciba por lo menos un pulso, obsérvese que el número de pulsos en ese intervalo tiene una distribución de Poisson con parámetro $\alpha t = 6(0.5) = 3$ Entonces con X = el número de pulsos recibidos en el intervalo de 30 segundos:

$$P(1 \leq X) = 1 - P(X = 0) = 1 - \frac{e^{-3}(3)^0}{0!} = 0.950$$

Contenidos

Contenidos

Distribución binomial negativa

La variable aleatoria y la distribución binomial negativa se basan en un experimento que satisface las siguientes condiciones:

1. El experimento consiste en una secuencia de ensayos independientes.
2. Cada ensayo puede dar por resultado un éxito (E) o una falla (F).
3. La probabilidad de éxito es constante de un ensayo a otro, por lo tanto $P(E \text{ en el ensayo } i) = p$ con $i = 1, 2, 3. \ldots$
4. El experimento continúa (se realizan ensayos) hasta que un total de r éxitos hayan sido observados, donde r es un entero positivo especificado.

La variable aleatoria de interés es X = el número de fallas que preceden al r-ésimo éxito; X se llama variable aleatoria binomial negativa porque, en contraste con la variable aleatoria binomial, el número de éxitos es fijo y el número de ensayos es aleatorio.

Proposición

La función masa de probabilidad de la variable aleatoria binomial negativa X con los parámetros r = número de éxitos (E) y p = P(E) es

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x \quad x = 0, 1, 2, \ldots$$

Contenidos

Distribución Geométrica

udp Escuela de Informática y Telecomunicaciones

La distribución geométrica es cualquiera de las dos distribuciones de probabilidad discretas siguientes:

- La distribución de probabilidad del número X del ensayo de Bernoulli necesaria para obtener un éxito, contenido en el conjunto 1, 2, 3,... o la distribución de probabilidad del número $Y = X - 1$ de fallos antes del primer éxito, contenido en el conjunto 0, 1, 2, 3,... .
- Si la probabilidad de éxito en cada ensayo es p, entonces la probabilidad de que x ensayos sean necesarios para obtener un éxito es $P(X = x) = (1 - p)^{x-1}p$; para x = 1, 2, 3,....

Continuación

- De forma análoga, la probabilidad de que haya x fallos antes del primer éxito es $P(Y = y) = p^y(1 - p)$ para y = 0, 1, 2,... . En ambos casos, la secuencia de es una progresión geométrica.
- El valor esperado y varianza de una variable aleatoria X distribuida geométricamente es: $E(X) = \frac{1}{p}$; $E(Y) = \frac{1-p}{p}$;$var(Y) = var(X) = \frac{1-p}{p^2}$

De una sala el 60 % de estudiantes son otakus, calcular probabilidad de extraer al 1er otaku, a la cuarta ocasión que extraemos un estudiante (con reposición).

Definir éxito: Que sea otaku (x=4;p=0.6;q=0.4).

$$P(X = 4) = (0.4)^{4-1}(0.6) = 0.04$$