DISTRIBUCIONES DE PROBABILIDAD MÁS USADAS

Resumen

Escrito por:

Prof. Arturo Rodrigo Farinha

Distribución Binomial

Situación

Sea un experimento en el cual solo puede suceder uno de dos resultados posibles (uno denominado *éxito* de probabilidad *p* y el otro denominado *fracaso* de probabilidad *1 - p*).
Sean *n* experimentos como ese, independientes entre sí.
Sea una variable aleatoria que cuenta la cantidad de éxitos sucedidos al hacer esos n experimentos.
Dicha VA tendrá una **distribución de probabilidad Binomial**.

Notación

$B(n, p)$

n cantidad de experimentos
p probabilidad de éxito en un experimento

Tipo de distribución

discreta

Probabilidad

La probabilidad de obtener exactamente k éxitos en n experimentos se calcula mediante la fórmula:

$$\boxed{P(x=k) = C_k^n . p^k . (1-p)^{n-k}}$$

$$C_k^n = \frac{n!}{(n-k)!k!}$$

Condiciones de aplicación:

$n \in N^+$

$p \in R / 0 < p < 1$

$k \in N / 0 \leq k \leq n$

Probabilidades Acumuladas

$$P(x \leq k) = \sum_{i=0}^{k} P(x=i)$$

$$P(x \geq k) = \sum_{i=k}^{n} P(x=i) = 1 - P(x \leq k-1)$$

Valores Destacados

Media (Esperanza, Valor Esperado)	np
Varianza	$np(1-p)$
Desviación Estándar	$\sqrt{np(1-p)}$
Sesgo *	$\frac{1-2p}{\sqrt{np(1-p)}}$
Curtosis *	$3 + \frac{1-6p(1-p)}{np(1-p)}$

* Explicación de lo que es Sesgo y Curtosis en pág. 7

Distribución Hipergeométrica

Consideraciones previas

La distribución *Binomial* se caracteriza por presentar experimentos de Bernoulli: situaciones dicotómicas (éxito – fracaso) y son considerados independientes entre sí debido a que las extracciones son:

con reposición o
sin reposición y la población es infinita (muy grande).

La distribución *Hipergeométrica* modeliza procesos de Bernoulli también dicotómicos, pero dependientes de los experimentos anteriores ya que no se reemplazan (devuelven) los objetos extraídos y la población es finita (pequeña).
Es decir que la distribución Hipergeométrica es especialmente útil en todos aquellos casos en los que se extraigan muestras o se realicen experiencias repetidas sin devolución del elemento extraído.
Es una distribución fundamental en el estudio de muestras pequeñas de poblaciones pequeñas y en el cálculo de probabilidades de juegos de azar y tiene grandes aplicaciones en el control de calidad en procesos experimentales en los cuales no es posible retornar a la situación de partida (los ensayos son destructivos).

Situación

Sea una población compuesta por **N** objetos de los cuales **r** cumplen cierta característica ("éxito").
De ella se extraen al azar y sin reposición **n** objetos.
Sea una variable aleatoria que cuenta la cantidad de objetos que cumplen la característica en dicha extracción.
Dicha VA tendrá una **distribución de probabilidad Hipergeométrica**.

Notación

$$Hip(N, n, r)$$

N número total de objetos
n número de objetos extraídos del total, sin devolverlos
r número de éxitos en el total de objetos

Tipo de distribución

discreta

Probabilidad

La probabilidad de obtener exactamente k éxitos en los n objetos extraídos se calcula mediante la fórmula:

$$\boxed{P(x=k)=\frac{C_k^r \,.\, C_{n-k}^{N-r}}{C_n^N}}$$

Condiciones de aplicación:

$$N \in N \,/\, N > 1$$

$$n \in N \,/\, 0 < n \leq N$$

$$r \in N \,/\, 0 < r \leq N$$

$$k \in N \,/\, Máx\{0, n-N+r\} \leq k \leq Mín\{n, r\}$$

Probabilidades Acumuladas

$$P(x \leq k) = \sum_{i=k_{mín}}^{k} P\left(x = i\right)$$

$$P(x \geq k) = \sum_{i=k}^{k_{máx}} P\left(x = i\right) = 1 - P(x \leq k - 1)$$

Valores Destacados

Media (Esperanza, Valor Esperado)	$\frac{n}{N} r$
Varianza	$\frac{nr(N-r)(N-n)}{N^2 (N-1)}$

Distribución de Poisson

Situación

Un experimento es de Poisson cuando que suceda algo (éxito) es expresado por unidad de tiempo, área, volumen, etc.
Sea una variable aleatoria que cuenta la cantidad de veces que sucede cierto evento.
Dicha VA tendrá una **distribución de probabilidad de Poisson**.

Notación

$Poi(\lambda)$ λ promedio de éxitos en el intervalo de interés

Tipo de distribución

discreta

Probabilidad

La probabilidad de obtener exactamente k ocurrencias en un experimento de Poisson se calcula mediante la fórmula:

$$\boxed{P(x=k)=\frac{e^{-\lambda}\lambda^{k}}{k!}}$$

Condiciones de aplicación:

$\lambda \in R^{+}$

$k \in N$

Probabilidades Acumuladas

$$P(x \le k)=\sum_{i=0}^{k} P(x=i)$$

$$P(x \ge k)=1-P(x \le k-1)$$

Valores Destacados

Media (Esperanza, Valor Esperado)	λ
Varianza	λ
Desviación Estándar	$\sqrt{\lambda}$
Sesgo *	$\frac{1}{\sqrt{\lambda}}$
Curtosis *	$3+\frac{1}{\lambda}$

* Explicación de lo que es Sesgo y Curtosis en pág. 7

Distribución Normal

Situación

Una variable aleatoria X tendrá una **distribución Normal** si su función de densidad tiene la siguiente forma:

$$\boxed{f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

Condiciones de aplicación :

$\mu\in R$

$\sigma\in R^+$

$x\in R$

Notación

$N\left(\mu,\sigma^2\right)$

μ media de la población

σ^2 varianza de la población

Tipo de distribución

continua

Probabilidades Acumuladas

$$P(x\le k)=\int_{-\infty}^{k}f(x)dx$$

$$P(x\ge k)=\int_{k}^{+\infty}f(x)dx=1-P(x\le k)$$

Valores Destacados

Media (Esperanza, Valor Esperado)	μ
Varianza	σ^2
Desviación Estándar	σ
Sesgo *	0
Curtosis *	3

* Explicación de lo que es Sesgo y Curtosis en pág. 7

¿Qué es el "sesgo"?

Es el grado de asimetría de la curva.

$$sesgo\begin{cases} >0 \;\Rightarrow\; \text{curva sesgada a la derecha} \\ <0 \;\Rightarrow\; \text{curva sesgada a la izquierda} \\ =0 \;\Rightarrow\; \text{curva simétrica} \end{cases}$$

¿Qué es la "curtosis"?

Expresa qué tan empinada se encuentra una distribución (cuánto más curtosis, más grande es el pico).

Se compara con la de la distribución Normal (3).

$$curtosis\begin{cases} <3 \;\Rightarrow\; \text{distribución menos "puntiaguda" que la dist. Normal} \\ >3 \;\Rightarrow\; \text{distribución más "puntiaguda" que la dist. Normal} \end{cases}$$

APROXIMACIONES

Aproximación de Poisson de la Binomial

$$Si \quad n \geq 30 \ y \ p \leq 0.1: \qquad B(n,p) \sim Poi(np)$$

Es decir que, si n y p cumplen las condiciones mencionadas, en lugar de la distribución *Binomial* se puede utilizar la distribución de *Poisson* con $\lambda = np$

Aproximación Binomial de la Hipergeométrica

$$Si \quad N \geq 50 \quad y \quad n \leq \frac{N}{10}: \qquad Hip(N,n,r) \sim B\left(n \ , \ \frac{r}{N}\right)$$

Es decir que, si N y n cumplen las condiciones mencionadas, en lugar de la distribución *Hipergeométrica* se puede utilizar la distribución *Binomial* con $n \quad y \quad p = \frac{r}{N}$

Aproximación Normal de la Binomial

$$Si \quad np \geq 5 \ y \ n(1-p) \geq 5: \qquad B(n,p) \sim N(np \ , \ np(1-p))$$

Es decir que, si n y p cumplen las condiciones mencionadas, en lugar de la distribución *Binomial* se puede utilizar la distribución *Normal* con $\mu = np$ y $\sigma = \sqrt{np(1-p)}$

En esta similitud no se debe olvidar que la Binomial es *discreta* y la Normal es *continua*. Debido a eso, hay que tener en cuenta un ajuste de media unidad (0.5) como se describe a continuación:

Sea X_B una variable aleatoria que presenta una distribución Binomial B(n , p).
Sea X_N una variable aleatoria que presenta una distribución Normal $N(np, np(1-p))$.

$$P(X_B = k) \cong P(k - 0.5 \leq X_N \leq k + 0.5)$$
$$P(X_B \leq k) \cong P(X_N \leq k + 0.5)$$
$$P(X_B \geq k) \cong P(X_N \geq k - 0.5)$$
$$P(k_1 \leq X_B \leq k_2) \cong P(k_1 - 0.5 \leq X_N \leq k_2 + 0.5)$$

Aproximación Normal de Poisson

$$\boxed{Si \quad \lambda \ es\ muy\ grande: \qquad Poi(\lambda) \sim N(\lambda\ , \lambda)}$$

Es decir que, si λ cumple la condición mencionada, en lugar de la distribución de *Poisson* se puede utilizar la distribución *Normal* con $\mu = \lambda$ y $\sigma = \sqrt{\lambda}$