Variables aleatorias

FaMAF

11 de marzo, 2010

Variables aleatorias discretas

Si X toma los valores $x_1, x_2, \ldots$:

- Función de masa de probabilidad:

$$p(x_i) = P(X = x_i).$$

$$\boxed{\sum_{i=1}^{\infty} p(x_i) = 1.}$$

$$P(X < b) = \lim_{n\to\infty} F(b - \frac{1}{n}).$$

Variables aleatorias continuas

Si X es una variable aleatoria continua:

$$P(X \in C) = \int_C f(x)\, dx.$$

- f es función de densidad.
- $F(a) = \int_{-\infty}^{a} f(x)\, dx.$
- $\frac{d}{da} F(a) = f(a).$
- $P(a - \epsilon/2 \leq X \leq a + \epsilon/2\} = \int_{a-\epsilon/2}^{a+\epsilon/2} f(x)\, dx \simeq f(a) \cdot \epsilon.$
- $P(X = a) = 0.$

Distribución conjunta

$$X,\ Y: \text{ variables aleatorias.}$$

- Función de distribución acumulada conjunta de X e Y.

$$F(a, b) = P(X \leq a,\ Y \leq b).$$

- X e Y discretas:

$$p(a, b) = P(X = a,\ Y = b)$$

- X e Y continuas son conjuntamente continuas si para todo C, D:

$$P(X \in C,\ Y \in D) = \int \int_{x \in C, y \in D} f(x, y)\, dx\, dy.$$

Distribuciones marginales

Si F es función de distribución conjunta de X e Y:

- $F_X(a) = P(X \leq a) = F(a, \infty)$: distribución marginal de X.

$$p_X(a) = \sum_b p(a, b).$$

$$F_X(a) = \int_{-\infty}^{a} \int_{-\infty}^{\infty} f(x, y)\, dy\, dx = \int_{-\infty}^{a} f_X(x)\, dx.$$

- $F_Y(b) = P(Y \leq b) = F(\infty, b)$: distribución marginal de Y.
- La distribución conjunta permite obtener las distribuciones (marginales) de X e Y.
- La recíproca no es cierta.

Distribución condicional

- X e Y discretas:

$$P(X = x \mid Y = y) = \frac{P(X = x \ \cap Y = y)}{P(Y = y)}.$$

$$p_{X|Y}(x \mid y) = \frac{p(x, y)}{p_Y(y)}.$$

- X e Y continuas:

$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x) f_X(x)}{f_Y(y)}.$$

$$\boxed{P(X \leq x \mid Y = y) = \int_{-\infty}^{x} f_{X|Y}(s \mid y)\, ds.}$$

Distribución conjunta

- X e Y son independientes si

$$P(X \in C,\ Y \in D) = P(X \in C) \cdot P(Y \in D).$$

  $\{X \in C\}$ y $\{Y \in D\}$ son eventos independientes.
- X e Y discretas:

$$P(X = a,\ Y = b) = P(X = a) \cdot P(Y = b)$$

$$p(a, b) = p_X(a) \cdot p_Y(b).$$

- X e Y continuas:

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

- Si X e Y son independientes, la distribución conjunta se obtiene a partir de las distribuciones de X e Y.

Valor esperado

- Si X es una v.a. discreta:

$$E[X] = \sum_i x_i \, P(X = x_i).$$

- Si X es una v.a. continua:

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

- El valor esperado no es necesariamente un valor posible de X.

Propiedades del valor esperado

Si $g : \mathbb{R} \mapsto \mathbb{R}$, entonces $g(X)$ es una v.a. y

- Discretas:
$$E[g(X)] = \sum g(x)\, p(x).$$
- Continuas:
$$E[g(X)] = \int g(x)\, f(x)\, dx.$$

- Linealidad:
$$E[aX + b] = aE[X] + b.$$

Si X e Y son v.a., entonces

$$E[X + Y] = E[X] + E[Y].$$

Varianza

Es una medida de la variación de X en torno a $\mu = E[X]$:

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

- No es lineal:
$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X).$$
- La varianza de la suma no es (necesariamente) la suma de las varianzas.
- Covarianza de X e Y:
$$\mathrm{cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)].$$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(X,Y)$.

Varianza y desviación estándar

- Si X e Y son independientes

$$\text{cov}(X, Y) = 0.$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

- Desviación estándar de X:

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

- Correlación de X e Y:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X) \cdot \sigma(Y)}.$$

Desigualdad de Chebyshev

- Si X toma sólo valores no negativos y $a > 0$, entonces

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Teorema (Desigualdad de Chebyshev)

Si X es v.a. con media μ y varianza σ^2, entonces para $k > 0$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Leyes de los grandes números

Si $X_1, X_2, \ldots, X_n, \ldots$ son v.a. independientes e idénticamente distribuidas, con media μ:

- Ley débil de los grandes números:

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) \to 0 \quad n \to \infty.$$

- Ley fuerte de los grandes números:
  Con probabilidad 1 se cumple que:

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu.$$

Uniforme

$$U\{1, n\}$$

n valores igualmente probables

- Rango: $\{1, \ldots, n\}$
- Función de masa:
$$P(X = i) = \frac{1}{n}.$$
- Valor esperado:
$$E[X] = \frac{n+1}{2}.$$
- Varianza:
$$\mathrm{Var}(X) = \frac{n^2 - 1}{12}.$$

Bernoulli

Éxito o fracaso

- Rango: $\{0, 1\}$
- Función de masa:

$$P(X = 1) = p, \qquad P(X = 0) = 1 - p.$$

- Valor esperado:

$$E[X] = p.$$

- Varianza:

$$\mathrm{Var}(X) = p \cdot (1 - p).$$

Binomial

$$Bi(n, p)$$

Número de éxitos en n ensayos independientes, con probabilidad p de éxito.

- Rango: $\{0, 1, 2, \ldots, n\}$
- Función de masa:

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

- Valor esperado:

$$E[X] = n\,p.$$

- Varianza:

$$\mathrm{Var}(X) = n\,p\,(1-p).$$

- Fórmula recursiva:

$$p_{i+1} = \frac{n-i}{i+1} \frac{p}{(1-p)} p_i.$$

Poisson

$$P(\lambda).$$

Número de éxitos en una cantidad grande de ensayos independientes

- Rango: $\{0, 1, 2, \dots\} = \{0\} \cup \mathbb{N}$
- Función de masa:
$$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}.$$
- Valor esperado:
$$E[X] = \lambda.$$
- Varianza:
$$\mathrm{Var}(X) = \lambda.$$
- Fórmula recursiva:
$$p_{i+1} = \frac{\lambda}{i+1} p_i.$$

Geométrica

$$\text{Geom}(p).$$

Número de ensayos independientes hasta el primer éxito

- Rango: $\mathbb{N}$.
- Función de masa:

$$P(X = n) = p\,(1-p)^{n-1}, \qquad n \geq 1.$$

- Valor esperado:

$$E[X] = \frac{1}{p}.$$

- Varianza:

$$\text{Var}(X) = \frac{1-p}{p^2}$$

Binomial Negativa o Pascal

$$\text{Bn}(r, p)$$

Número de ensayos independientes hasta obtener r éxitos

- Rango: $\{r, r+1, r+2, \dots\} = \{n \in \mathbb{N} \mid n \geq r\}$.
- Función de masa:

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \qquad n \geq r.$$

- Valor esperado:

$$E[X] = \frac{r}{p}.$$

- Varianza:

$$\text{Var}(X) = \frac{r(1-p)}{p^2}.$$

Hipergeométrica

Número de éxitos en una muestra de tamaño n extraída de un conjunto de $N+M$ elementos

- Rango: $\{0, 1, 2, \dots, n\}$
- Función de masa:

$$P(X = i) = \frac{\binom{N}{i}\binom{M}{n-i}}{\binom{N+M}{n}}.$$

- Valor esperado:

$$E[X] = \frac{nN}{N+M}.$$

- Varianza:

$$\mathrm{Var}(X) = \frac{nNM}{(N+M)^2}\left(1 - \frac{n-1}{N+M-1}\right).$$