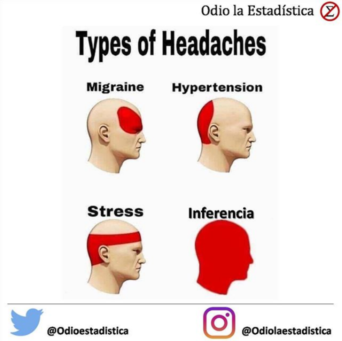
Odio la Estadística
Types of Headaches
Migraine
Hypertension
Stress
Inferencia
@Odioestadistica
@Odiolaestadistica

Probabilidad - Clase 26
Varianza de una variable aleatoria

Ernesto Mordecki

Facultad de Ciencias, Universidad de la República. Montevideo, Uruguay

Curso de la Licenciatura en Matemática - 2020

Contenidos

Varinza

- Queremos medir cuanto se diferencia una v.a. de su valor esperado
- Sabemos que si $\mathbf{P}(X = c) = 1$ (una constante), entonces $\mathbf{E}\, X = c$.
- Pero si X toma valores distribuídos, ¿cuánto se *dispersan* esos valores?
- Para eso definimos la **varianza** de una variable aleatoria, que notamos $\mathbf{var}\, X$
- Su **desvío estándar**, notado $\sigma_X := \sqrt{\mathbf{var}\, X}$

Definición de varianza

- Consideremos un espacio $(\Omega, \mathcal{A}, \mathbf{P})$
- Consideremos una variable aleatoria $X = X(\omega),\ \omega \in \Omega$.
- Supongamos que existe la esperanza matemática $\mathbf{E}\, X = a$.
- Definimos (cuando existe) la *varianza* de X, como la esperanza matemática

$$\mathbf{var}\, X := \mathbf{E}(X - a)^2,$$

- Al valor positivo de la raíz cuadrada de la varianza le llamamos *desviación estándar* de la variable aleatoria X:

$$\sigma_X := \sqrt{\mathbf{var}\, X},$$

Con la integral de Lebesgue (que discutimos la clase pasada) o la de Riemann-Stieltjes en el caso general se obtienen las fórmulas

$$\mathbf{var}\, X = \int_\Omega (X-a)^2 d\,\mathbf{P} = \int_{-\infty}^{\infty} (x-a)^2 dF(x), \tag{1}$$

y la varianza existe si alguna de las integrales anteriores es convergente (lo que equivale, en este caso, a la convergencia absoluta), donde suponemos que existe la esperanza

$$a = \mathbf{E}\, X = \int_\Omega X d\,\mathbf{P} = \int_{-\infty}^{\infty} x dF(x).$$

Caso discreto

Si la variable aleatoria X tiene distribución discreta, y toma los valores $x_1, x_2, \ldots$ con probabilidades $p_1, p_2, \ldots$ respectivamente, entonces

$$\mathbf{var}\, X = \sum_k (x_k - a)^2 p_k, \tag{2}$$

si la serie anterior es convergente, donde suponemos que existe la esperanza

$$a = \sum_k x_k p_k.$$

Caso continuo

Si X es una variable aleatoria con densidad $p(x)$, entonces

$$\textbf{var}\, X = \int_{-\infty}^{\infty} (x-a)^2 p(x) dx, \tag{3}$$

si la integral anterior es convergente, donde suponemos que existe la esperanza

$$a = \int_{-\infty}^{\infty} x p(x) dx.$$

Primeras propiedades

Veamos dos identidades útiles, que verifica la varianza de una variable aleatoria arbitraria X. Tenemos

$$\mathbf{var}\, X = \mathbf{E}(X^2) - (\mathbf{E}\, X)^2, \tag{4}$$

$$\mathbf{var}\, X = \mathbf{E}\big(X(X-1)\big) - \mathbf{E}\, X(\mathbf{E}\, X - 1). \tag{5}$$

Demostremos (4). Aplicando la definición de varianza

$$\begin{aligned}\mathbf{var}\, X &= \mathbf{E}(X^2 - 2aX + a^2) = \mathbf{E}(X^2) - 2a\,\mathbf{E}\, X + a^2 \\ &= \mathbf{E}(X^2) - (\mathbf{E}\, X)^2.\end{aligned}$$

Para demostrar (5), tenemos

$$\mathbf{E}\big(X(X-1)\big) - \mathbf{E}\,X(\mathbf{E}\,X - 1) = \mathbf{E}(X^2) - \mathbf{E}\,X - (\mathbf{E}\,X)^2 + \mathbf{E}\,X$$
$$= \mathbf{E}(X^2) - (\mathbf{E}\,X)^2 = \mathbf{var}\,X.$$

Ejemplo Calcular la varianza de una variable aleatoria X con distribución binomial con parámetros (n, p).
Sabemos que $\mathbf{E}\,X = np$. Utilicemos la identidad (5), especialmente adecuada si X toma únicamente valores enteros. Si $q = 1 - p$, tenemos

$$\mathbf{E}\big(X(X-1)\big) = \sum_{k=0}^{n} k(k-1)\,\mathbf{P}(X=k) = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k}$$
$$= n(n-1)p^2 \sum_{i=0}^{n-2} \frac{(n-2)!}{i!(n-2-i)!} p^i q^{n-2-i}$$
$$= n(n-1)p^2 (q+p)^{n-2} = n(n-1)p^2.$$

Por ésto, de (5) obtenemos

$$\mathbf{var}\,X = n(n-1)p^2 - np(np-1) = np(1-p) = npq.$$

Ejemplo Calcular la varianza de una variable aleatoria X con distribución de Poisson con parámetro $\lambda > 0$.
Tenemos

$$\mathbf{E}\big(X(X-1)\big) = \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$$
$$= \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2.$$

Por ésto, teniendo en cuenta que $\mathbf{E}\,X = \lambda$ (ver ejemplo **??**) y la fórmula (5), obtenemos

$$\mathbf{var}\,X = \mathbf{E}\big(X(X-1)\big) - \mathbf{E}\,X(\mathbf{E}\,X - 1) = \lambda^2 - \lambda(\lambda - 1) = \lambda.$$

Variable normal (a, σ^2)

Calcular la varianza de una variable aleatoria X con distribución normal con parámetros (a, σ).
Sabemos $\mathbf{E}\, X = a$. Aplicando (3), tenemos

$$\mathbf{var}\, X = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-(x-a)^2/(2\sigma^2)} dx.$$

Introducimos el cambio de variable $u = (x-a)/\sigma$ y aplicamos la fórmula de integración por partes, para obtener

$$\mathbf{var}\, X = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-u^2/2} du = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u d\Big(- e^{-u^2/2}\Big)$$
$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} du = \sigma^2.$$

- De esta forma **var** $X = \sigma^2$, y en consecuencia σ es la desviación estándar de la variable aleatoria X.
- Encontramos así el sentido probabilístico de los parámetros a y σ que caracterizan a una distribución normal.

Propiedad. Para cualquier variable aleatoria X se verifica **var** $X \geq 0$.

Demostración: Esta conclusión es inmediata a partir de cualquiera de las definiciones que se utilice.

Varianza nula

Propiedad. Se verifica $\mathbf{var}\, X = 0$ si y solo si la variable aleatoria X tiene distribución degenerada.

Demostración Supongamos que X tiene distribución degenerada, es decir $\mathbf{P}(X = c) = 1$ para alguna constante c real. Entonces $\mathbf{E}\, X = c$, y
$\mathbf{var}\, X = \mathbf{E}(X - c)^2 = (c - c)^2 \times 1 = 0.$

Supongamos ahora que $\mathbf{var}\, X = 0$ y sea $a = \mathbf{E}\, X$. Demostremos primero que X tiene distribución degenerada cuando tiene distribución discreta, y toma los valores $x_1, x_2, \dots$ con probabilidades $p_1, p_2, \dots$ respectivamente. En este caso tenemos

$$\mathbf{var}\, X = \sum_k (x_k - a)^2 p_k = 0.$$

Entonces cada sumando es nulo, y $x_k = a$ para todo $k = 1, 2, \dots$ (porque $p_k > 0$). Por ésto $\mathbf{P}(X = a) = p_1 + p_2 + \cdots = 1$, y X tiene distribución degenerada.

Propiedad. Sea X una variable aleatoria con varianza $\mathbf{var}\, X$, y sean a, b constantes. Entonces

$$\mathbf{var}(aX + b) = a^2\, \mathbf{var}\, X.$$

Demostración Por la definición de varianza, tenemos

$$\begin{aligned}\mathbf{var}(aX + b) &= \mathbf{E}\big(aX + b - \mathbf{E}(aX + b)\big)^2 = \mathbf{E}\big(aX + b - a\,\mathbf{E}\, X - b\big)^2 \\ &= \mathbf{E}\big(a(X - \mathbf{E}\, X)\big)^2 = a^2\, \mathbf{E}(X - \mathbf{E}\, X)^2 = a^2\, \mathbf{var}\, X,\end{aligned}$$

concluyendo la demostración.

Propiedad. Sean X e Y variables aleatorias independientes[1], para las cuales existe la varianza. Entonces

$$\mathbf{var}(X+Y) = \mathbf{var}\,X + \mathbf{var}\,Y.$$

Esta propiedad es un caso particular de la siguiente.

Propiedad. Sean $X_1, \ldots, X_n$ variables aleatorias independientes dos a dos, para las cuales existe la varianza. Entonces

$$\mathbf{var}(X_1 + \cdots + X_n) = \mathbf{var}\,X_1 + \cdots + \mathbf{var}\,X_n. \tag{6}$$

[1] OjO

Demostración Designemos $a_k = \mathbf{E}\, X_k$ $(k = 1, \dots, n)$.
Partiendo de la definición de varianza, tenemos

$$\begin{aligned}
\mathbf{var} \sum_{k=1}^{n} X_k &= \mathbf{E}\Big(\sum_{k=1}^{n} X_k - \mathbf{E} \sum_{k=1}^{n} X_k\Big)^2 = \mathbf{E}\Big(\sum_{k=1}^{n} (X_k - a_k)\Big)^2 \\
&= \sum_{k=1}^{n} \mathbf{E}(X_k - a_k)^2 + 2 \sum_{1 \le k < j \le n} \mathbf{E}(X_k - a_k)(X_j - a_j) \\
&= \sum_{k=1}^{n} \mathbf{var}\, X_k,
\end{aligned}$$

porque

$$\mathbf{E}(X_k - a_k)(X_j - a_j) = \mathbf{E}(X_k - a_k)\, \mathbf{E}(X_j - a_j) = 0.$$

Desigualdad de Chebishev

Lema

- *Consideremos una variable aleatoria X no negativa,*
- *supongamos que existe la esperanza matemática $\mathbf{E}\,X$.*
- *Entonces, para cualquier $t > 0$, tenemos*

$$\mathbf{P}(X \geq t) \leq \frac{1}{t}\,\mathbf{E}\,X. \tag{7}$$

Demostración

- Demostremos primero este lema cuando X tiene distribución discreta,
- toma los valores $x_1, x_2, \ldots,$ con probabilidades $p_1, p_2, \ldots,$ respectivamente.
- Como $x_k \geq 0$ para todo k, tenemos

$$\mathbf{E}\, X = \sum_k x_k p_k \geq \sum_{k\colon\, x_k \geq t} x_k p_k \geq \sum_{k\colon\, x_k \geq t} t p_k = t\, \mathbf{P}(X \geq t),$$

y de aquí se obtiene (7).

Caso continuo

- Si X tiene densidad $p(x)$ resulta que $p(x) = 0$ para todo $x < 0$, porque $x < 0$ implica $F(x) = \mathbf{P}(X \leq x) = 0$ cuando $X \geq 0$.
- Por ésto,

$$\begin{aligned}
\mathbf{E}\, X &= \int_{-\infty}^{\infty} xp(x)dx = \int_{0}^{\infty} xp(x)dx \\
&\geq \int_{t}^{\infty} xp(x)dx \geq \int_{t}^{\infty} tp(x)dx \\
&= t \int_{t}^{\infty} p(x)dx = t\,\mathbf{P}(X \geq t),
\end{aligned}$$

obteniendo (7).

- Este lema fue demostrado por Pafnuty Chebyshev (Chebishev: 1821-1894).
- La condición $X \geq 0$ no se puede omitir;
- sin ella podría ocurrir que $\mathbf{E}\, X < 0$, lo que es incompatible con la desigualdad (7).

Sea ahora X una variable aleatoria arbitraria, para la cual existe la varianza. Designemos $Z = (X - \mathbf{E}\,X)^2$. Es claro que Z toma valores no negativos, y que $\mathbf{E}\,Z = \mathbf{var}\,X$. Por el lema anterior tenemos que para todo $\varepsilon > 0$, se verifica

$$\mathbf{P}\big((X - \mathbf{E}\,X)^2 \geq \varepsilon^2\big) \leq \frac{1}{\varepsilon^2}\,\mathbf{E}\,Z = \frac{1}{\varepsilon^2}\,\mathbf{var}\,X,$$

que, tomando raíz cuadrada, es

$$\mathbf{P}\big(|X - \mathbf{E}\,X| \geq \varepsilon\big) \leq \frac{1}{\varepsilon^2}\,\mathbf{var}\,X. \tag{8}$$

La desigualdad (8), válida para cualquier variable con varianza y para cualquier $\varepsilon > 0$, se denomina *desigualdad de Chebishev*.

Esta desigualdad resulta ser una de las más utilizadas en la teoría de la probabilidad. Teniendo en cuenta, que la suma de las probabilidades de un suceso y su contrario es igual a uno, la desigualdad de Chebishev (8) se puede escribir de la forma

$$\mathbf{P}\big(|X - \mathbf{E}\,X| < \varepsilon\big) \geq 1 - \frac{1}{\varepsilon^2}\,\mathbf{var}\,X,$$

para cualquier $\varepsilon > 0$.


DIME ALBERT ¿CREES QUE SOY BONITA?
¿BONITA COMPARADA CON QUIÉN?
Daniel PAZ
Einstein a punto de descubrir que la relatividad mal aplicada puede costarle un paraguazo