Fontanarrosa

Figura: Inodoro Pereyra - el renegau

Probabilidad - Clase 28
Desigualdad de Chebishev y otros momentos

Ernesto Mordecki

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

Curso de la Licenciatura en Matemática - 2020

Contenidos

Varianza

- Queremos medir cuanto se diferencia una v.a. de su valor esperado
- Para eso definimos la **varianza** de una variable aleatoria, que notamos $\mathbf{var}\, X$
- Su **desvío estándar**, notado $\sigma_X := \sqrt{\mathbf{var}\, X}$
- Supongamos que existe la esperanza matemática $\mathbf{E}\, X = a$.
- Definimos (cuando existe) la *varianza* de X, como la esperanza matemática

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

- La *desviación estándar* es

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

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

$$\mathbf{var}(X_1 + \cdots + X_n) = \mathbf{var}\, X_1 + \cdots + \mathbf{var}\, X_n. \tag{1}$$

[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*

$$\mathbf{P}(X \geq t) \leq \frac{1}{t}\,\mathbf{E}\,X. \tag{2}$$

Demostración

- Demostremos primero este lema cuando X tiene distribución discreta,
- toma los valores $x_1, x_2, \dots,$ con probabilidades $p_1, p_2, \dots,$ 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 (2).

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 (2).

- 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 (2).

- Sea X una variable con varianza.
- Sea

$$Z = (X - \mathbf{E}\,X)^2 \geq 0, \text{ que nos da } \mathbf{E}\,Z = \mathbf{var}\,X.$$

- Para $\varepsilon > 0$,

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

- Tomando raíz cuadrada, es

$$\mathbf{P}\big(|X - \mathbf{E}\,X| \geq \varepsilon\big) \leq \frac{1}{\varepsilon^2}\,\mathbf{var}\,X. \tag{3}$$

La desigualdad (3), 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.
- Tomando complemento

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

  para cualquier $\varepsilon > 0$.
- Si μ es Binomial con parámetros (n, p), tenemos $\mathbf{E}\,\mu = np$, $\mathbf{var}\,\mu = npq$, la desigualdad de Chebishev es

$$\mathbf{P}(|\mu - np| \geq n\varepsilon) \leq \frac{npq}{n^2\varepsilon^2}$$

$$\mathbf{P}\left(\left|\frac{\mu}{n} - p\right| \geq \varepsilon\right) \leq \frac{pq}{n\varepsilon^2}$$

  (¡nuestra vieja conocida!)

Otros momentos

- Definimos el *momento de orden* k de la variable aleatoria X:
$$\alpha_k = \mathbf{E}(X^k),$$
cuando la esperanza anterior existe. En particular, $\alpha_1 = \mathbf{E}\, X$.

- Observemos que la existencia del momento de orden k de una variable aleatoria es equivalente a la finitud del así llamado *momento absoluto de orden* k,
$$\beta_k = \mathbf{E}\, |X|^k.$$

- Si $|X| \leq C$ (X es acotada) existen α_k y β_k para cualquier k.

Proposición

Sea X una variable aleatoria. Si existe su momento de orden k, entonces, existe su momento de orden m, para todo $m = 1, \ldots, k$.

Demostración Elegimos $1 \leq m \leq k$. Vale la desigualdad $|X|^m \leq 1 + |X|^k$. En efecto, si $|X| \leq 1$, se tiene $|X|^m \leq 1$, y si $|X| \geq 1$, se tiene $|X|^m \leq |X|^k$. De esta desigualdad se obtiene que $\mathbf{E}\,|X|^m \leq 1 + \mathbf{E}\,|X|^k$, es decir, $\beta_m \leq 1 + \beta_k$ si $1 \leq m \leq k$. Como β_k es finito, β_m es finito, concluyendo la demostración.

- Definimos el *momento centrado de orden k* de X,

$$\mu_k = \mathbf{E}(X - \mathbf{E}\,X)^k,$$

cuando la esperanzas anteriores existen.

- Como ocurre con el momento de orden k, la existencia del momento centrado de orden k de una variable aleatoria es equivalente a la finitud del *momento absoluto centrado de orden k*, designado ν_k y definido mediante la identidad

$$\nu_k = \mathbf{E}\,|X - \mathbf{E}\,X|^k.$$

cuando la esperanza anterior es finita.

- En particular, $\mu_2 = \nu_2 = \mathbf{var}\,X$.

- Definimos el *momento absoluto inicial* y el *momento absoluto centrado* de orden p de una variable aleatoria X, mediante las fórmulas

$$\beta_p = \mathbf{E}\,|X|^p, \qquad \nu_p = \mathbf{E}\,|X - \mathbf{E}\,X|^p,$$

respectivamente, si existen las esperanzas.

- De la desigualdad de Cauchy-Bunyakovsky se obtiene, que $\mathbf{E}\,|X| \leq \sqrt{\mathbf{E}(X^2)}$, es decir $\beta_1 \leq \beta_2^{1/2}$, si los momentos anteriores existen. Esta desigualdad se obtiene también de la desigualdad de Lyapunov: $\beta_s^{1/s} \leq \beta_r^{1/r}$, si $0 < s < r$. (Estas desigualdades se presentan sin aquí demostración)

- La *función generatriz de momentos* de una variable aleatoria X, mediante la identidad

$$M(t) = \mathbf{E}\, e^{tX},$$

para los t que dan finito.

- Si esta función está definida para todo t de un intervalo $|t| \leq b$, en este intervalo vale la fórmula

$$M(t) = 1 + \sum_{k=1}^{\infty} \alpha_k \frac{t^k}{k!},$$

que surge de su desarrollo de Taylor.

- Si $|X| \leq C$ para alguna constante C, entonces, la función generatriz de momentos está definida para todo t real.

- SI X es normal, también

Cuantiles, percentiles

- Como los momentos de una variable aleatoria no siempre existen, se introducen otras características numéricas de las variables aleatorias, que existen siempre.
- Consideremos una variable aleatoria X y un número $0 < q < 1$.
- Llamamos *cuantil* de orden q de la variable aleatoria X, a cualquier número κ_q que verifique las condiciones
$$\mathbf{P}(X \leq \kappa_q) \geq q, \qquad \mathbf{P}(X \geq \kappa_q) \geq 1 - q.$$
- Si la función de distribución de una variable aleatoria X es estrictamente creciente en toda la recta real, su cuantil de cualquier orden q es único.
- Una variable aleatoria arbitraria no siempre cumple esta condición.

- El cuantil de orden $1/2$ se denomina *mediana*. De esta forma, la mediana de una variable aleatoria X es cualquier número m, que verifica las condiciones

$$\mathbf{P}(X \leq m) \geq \frac{1}{2}, \qquad \mathbf{P}(X \geq m) \geq \frac{1}{2}.$$

- Los cuantiles de orden $1/4$, $1/2$ y $3/4$, se denominan *cuartiles*.

Moda

- La *moda* de una variable aleatoria X se define únicamente para variables aleatorias discretas continuas.

- Si X toma los valores $x_1 < x_2 < x_3 < \cdots$, con probabilidades $p_1, p_2, p_3, \ldots$, entonces, la *moda* de X es cualquier número x_k para el que se cumpla $p_{k-1} \leq p_k \geq p_{k+1}$.

- Si X tiene densidad $p(x)$, la *moda* de X es cualquier punto de máximo local de $p(x)$.

- Es claro, que en general, la moda no es única.

- Si una variable aleatoria tiene una única moda su distribución se denomina *unimodal*.

- $X \sim \mathcal{N}(a, \sigma^2)$, la esperanza, la mediana y la moda (únicas) coinciden con a.

- ¿Sabes inglés?
- Claro
- ¿Que significa never?
- Nunca
- ¿Y never ever?
- Nunca unca