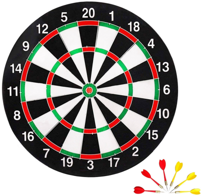
Dos estadísticos disparan a un blanco:



El primero erra por un metro a la derecha, el segundo, un metro a la izquierda. Luego, chocan los cinco con expresión de alegría. Alguien que pasa por allí y no entiende nada. Les pregunta:

- ¿por que festejan? erraron los dos por mucho
- son formas de verlo, responde uno;
- en promedio: ¡acertamos! dice el otro.

Probabilidad - Clase 29
Ley de los grandes números

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 $X_1, \dots, X_n$ 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

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

- 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

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

Es la *Desigualdad de Chebishev*.

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 $0 < q < 1$.

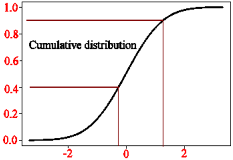
- Llamamos *cuantil* de orden q de la variable aleatoria X, a cualquier número κ_q que verifique las condiciones

$$\mathbf{P}(X \leq \kappa_q) \geq q, \qquad \mathbf{P}(X \geq \kappa_q) \geq 1 - q.$$

- 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.

Mediana



- 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

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.

Ley de los grandes números

- La sucesión de variables aleatorias $Y_1, Y_2 \ldots$ converge en probabilidad a la variable Y su para todo $\varepsilon > 0$ tenemos

$$\mathbf{P}(|Y_n - Y| > \varepsilon) \to 0 \quad (n \to \infty).$$

- Escribimos

$$Y_n \xrightarrow{\mathbf{P}} Y.$$

- Muchas veces Y es una constante.

- Tenemos una *ley de los grandes números* cuando los promedios aritméticos de una cantidad creciente a infinito de variables aleatorias convergen en probabilidad. [2]

---

[2] Si tenemos convergencia en probabilidad, decimos *ley débil de los grandes números*, si tenemos convergencia con probabilidad 1, decimos *ley fuerte de los grandes números*.

Teorema (Chebishev)

Consideremos una sucesión de variables aleatorias

$$X_1, X_2, \ldots,$$

independientes dos a dos, y con esperanzas matemáticas $a_1, a_2, \ldots$. *Supongamos que se cumple la condicion*

$$\textbf{var}\, X_n \leq C, \qquad n = 1, 2, \ldots,$$

donde C *es una constante arbitraria. Entonces*

$$\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} a_i \xrightarrow{\mathbf{P}} 0. \tag{3}$$

Variables con esperanza constante

- Es importante el caso particular en el que las esperanzas son iguales,

$$\mathbf{E}\, X_n = a, \qquad \text{para todo } n = 1, 2, \ldots .$$

- Tenemos

$$\frac{1}{n}\sum_{i=1}^{n} a_i = a,$$

- La convergencia en (3) se transforma en

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbf{P}} a.$$

- Si las variables son i.i.d, entonces $\mathbf{E}\, X_n = a$ para todo n.

Demostración

- Consideremos para cada $n = 1, 2, \ldots$

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

- Es claro que

$$\mathbf{E}\, Z_n = \frac{1}{n} \sum_{i=1}^{n} a_i.$$

- Como las variables aleatorias son independientes dos a dos,

$$\mathbf{var} \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \mathbf{var}\, X_i \leq nC$$

en vista de la propiedad anterior.

- Aplicamos ahora la desigualdad de Chebishev:

$$\mathbf{P}\left(\Big|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}a_i\Big|\geq\varepsilon\right)=\mathbf{P}\big(|Z_n-\mathbf{E}\,Z_n|\geq\varepsilon\big)$$
$$\leq\frac{1}{\varepsilon^2}\,\mathbf{var}\,Z_n=\frac{1}{n^2\varepsilon^2}\,\mathbf{var}\sum_{i=1}^{n}X_i\to 0,$$

si $n\to\infty$, dado que $\sum_{i=1}^{n}\mathbf{var}\,X_i\leq nC$.
Esto concluye la demostración.

- El teorema de Chebishev da fundamento a la *regla de la media aritmética*:
- para la estimación del valor de una constante física a, desconocida, mediante los resultados de n mediciones de su magnitud, se recomienda tomar el promedio aritmético de estas mediciones.
- Veamos la fundamentación: $X_1, \ldots, X_n$ son los resultados de las n mediciones de esta constante a.
- Como las mediciones tienen errores[3], consideramos

$$X_i = a + \Delta_i,$$

donde Δ_i es el error que se comete en la i–ésima medición.

[3] ¡Caso contrario sería suficiente una única medición!

Suponemos que

$$\mathbf{E}\,\Delta_i = 0, \qquad \mathbf{var}\,\Delta_i = \sigma^2 \qquad (i = 1, \dots, n), \tag{4}$$

donde el valor de σ^2 puede ser desconocido. Las condiciones en (4) son equivalentes a la condiciones

$$\mathbf{E}\,X_i = a, \qquad \mathbf{var}\,X_i = \sigma^2 \qquad (i = 1, \dots, n). \tag{5}$$

- $\mathbf{E}\,X_i = a$ se interpreta como la ausencia de error sistemático en las mediciones,
- $\mathbf{var}\,X_i = cte$ significa que las mediciones tienen la misma precisión.
- Si se supone además que $X_1, \dots X_n$ son independientes dos a dos, el teorema de Chebishev con $a_1 = \cdots = a_n = a$ da

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\mathbf{P}} a.$$

En consecuencia:

- el promedio aritmético de los resultados de las mediciones de una constante física a converge en probabilidad al valor de esta constante;
- en la terminología de la estadística, el promedio aritmético de las mediciones es un *estimador consistente* de la constante desconocida a.

Corolario: teorema de Bernoulli

- Consideremos una serie de n experimentos independientes, con dos resultados posibles cada uno (éxito y fracaso), y probabilidad de éxito igual a p en cada experimento ($0 < p < 1$).
- Sea μ la cantidad de éxitos en n experimentos.
- Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias, cada una de las cuales toma el valor 1 con probabilidad p (si ocurre un éxito) y el valor 0 con probabilidad $q = 1 - p$ (si ocurre un fracaso).

- Tenemos $\mathbf{E}\, X_i = p$, $\mathbf{var}\, X_i = pq \leq 1$ para cada $i = 1, 2, \ldots$.
- Además $\mu = \sum_{i=i}^{n} X_i$, porque la suma contiene tantos sumandos iguales a uno como éxitos ocurren en los primeros n experimentos, siendo nulos los sumandos restantes.
- Las variables aleatorias $X_1, X_2, \ldots$ son independientes dos a dos y cumplen la condición $\mathbf{var}\, X_i \leq 1$ $(i = 1, 2, \ldots)$, por lo que es aplicable el teorema de Chebishev:

$$\mu/n \xrightarrow{\mathbf{P}} p.$$

- La última afirmación significa que la frecuencia de éxitos en n experimentos, converge en probabilidad a p, la probabilidad de éxito en un experimento.

Chiste tonto (sic)