# Probabilidad y Estadística

Agustín G. Bonifacio

UNSL

## Esperanza Matemática y Varianza

## Definición

Sea $X$ una v.a. con función de densidad $f_X$. La esperanza matemática (o media) de $X$, $E(X)$, se define por:

1. $$E(X) = \sum_{\{r_i \in RgX\}} r_i f_X(r_i)$$

   si $X$ es discreta y la serie involucrada en la definición es absolutamente convergente (esto es, si $\sum_{\{r_i \in RgX\}} |r_i| f_X(r_i) < \infty$),

2. $$E(X) = \int_{-\infty}^{\infty} t f_X(t) dt$$

   si $X$ es continua y la integral involucrada en la definición es absolutamente convergente (esto es, si $\int_{-\infty}^{\infty} |t| f_X(t) dt < \infty$)

## Observaciones

- La *esperanza matemática* de una v.a. puede interpretarse como un promedio 'pesado' de los valores que la variable puede tomar.
- La *esperanza matemática* también puede mirarse como el promedio aritmético de los valores que aparecen en una muestra 'grande' de la variable aleatoria.

Veamos que, para $N$ grande, $\bar{x} \approx E(X)$ en el caso de que $X$ es v.a. discreta con rango finito.

## Esperanza matemática de algunas v.a. discretas

- Si $X \sim$ Bernoulli$(p)$, entonces $E(X) = p$.
- Si $X \sim$ Unif$(n)$, entonces $E(X) = \frac{n+1}{2}$.
- Si $X \sim$ Binom$(n, p)$, entonces $E(X) = np$.
- Si $X \sim$ Poisson$(\lambda)$, entonces $E(X) = \lambda$.
- Si $X \sim$ Geom$(p)$, entonces $E(X) = \frac{1}{p}$.

## Esperanza matemática de algunas v.a. continuas

- Si $X \sim$ Unif$[a, b]$, entonces $E(X) = \frac{b+a}{2}$.
- Si $X \sim N(\mu, \sigma^2)$, entonces $E(X) = \mu$.
- Si $X \sim$ Exp$(\lambda)$, entonces $E(X) = \frac{1}{\lambda}$.

## Propiedades

Sean $X, Y$ v.a. y sea $c \in \mathbb{R}$. Entonces:

1. $E(X + c) = E(X) + c$,
2. $E(cX) = cE(X)$,
3. $E(X + Y) = E(X) + E(Y)$.

- Vimos que la *esperanza* $E(X)$ puede interpretarse como un promedio ponderado de los valores que la v.a. $X$ puede tomar.
- La *esperanza* es una medida de centralización que indica alrededor de qué valor tienden a ubicarse los valores de $X$ si se toma una muestra grande para $X$.
- Sin embargo, v.a. muy diferentes pueden tener la misma esperanza.

**Ejemplo**

1. Sea la v.a. $X$ tal que $RgX = \{0, 1\}$, $f_X(0) = 1$ y $f_X(1) = 0$.
2. Sea la v.a. $Y$ tal que $RgY = \{-1, 1\}$, $f_Y(-1) = 1/2$ y $f_Y(1) = 1/2$.
3. Sea la v.a. $Z$ tal que $RgZ = \{-100, 100\}$, $f_Z(-100) = 1/2$ y $f_Z(100) = 1/2$.

Si calculamos las esperanzas obtenemos $E(X) = E(Y) = E(Z) = 0$. Sin embargo, la dispersión de los datos alrededor de la media aumenta de $X$ a $Y$ y de $Y$ a $Z$. El concepto de *varianza* es una medida de esa dispersión.

## Definición

Sea $X$ una v.a. con función de densidad $f_X$. La varianza de $X$, $Var(X)$, se define por:

1. $$Var(X) = \sum_{\{r_i \in RgX\}} [r_i - E(X)]^2 f_X(r_i)$$

   si $X$ es discreta y la serie involucrada en la definición es absolutamente convergente,

2. $$Var(X) = \int_{-\infty}^{\infty} [t - E(X)]^2 f_X(t) dt$$

   si $X$ es continua y la integral involucrada en la definición es absolutamente convergente.

## Observaciones

- $Var(X) \geq 0$,
- La raíz cuadrada de la *varianza* se denomina desviación típica (o estándar) de $X$.

## Ejemplo

1. Sea la v.a. $X$ tal que $RgX = \{0, 1\}$, $f_X(0) = 1$ y $f_X(1) = 0$.
2. Sea la v.a. $Y$ tal que $RgY = \{-1, 1\}$, $f_Y(-1) = 1/2$ y $f_Y(1) = 1/2$.
3. Sea la v.a. $Z$ tal que $RgZ = \{-100, 100\}$, $f_Z(-100) = 1/2$ y $f_Z(100) = 1/2$.

Veamos que las varianzas de $X, Y$ y $Z$ son diferentes. De hecho, $Var(X) = 0, Var(Y) = 1$ y $Var(Z) = 10{,}000$.

- A continuación damos una fórmula alternativa de cálculo para la *varianza.*
- Sea

$$E(X^2) = \sum_{\{r_i \in RgX\}} r_i^2 f_X(r_i)$$

si $X$ es discreta y

$$E(X^2) = \int_{-\infty}^{\infty} t^2 f_X(t) dt$$

si $X$ es continua.

**Teorema**

Sea $X$ una v.a. tal que $E(X^2) < \infty$. Entonces

$$Var(X) = E(X^2) - [E(X)]^2.$$

## Propiedades

Sean $X$ una v.a. y $c \in \mathbb{R}$. Entonces:

1. $Var(c) = 0$,
2. $Var(X + c) = Var(X)$,
3. $Var(cX) = c^2 Var(X)$,
4. $Var(X) = E(X^2) - [E(X)]^2$.

## Varianza de algunas v.a. discretas

- Si $X \sim$ Bernoulli$(p)$, entonces $Var(X) = p(1-p)$.
- Si $X \sim$ Unif$(n)$, entonces $Var(X) = (n^2-1)/12$.
- Si $X \sim$ Binom$(n,p)$, entonces $Var(X) = np(1-p)$.
- Si $X \sim$ Poisson$(\lambda)$, entonces $Var(X) = \lambda$.
- Si $X \sim$ Geom$(p)$, entonces $Var(X) = (1-p)/p^2$.

## Varianza de algunas v.a. continuas

- Si $X \sim$ Unif$[a,b]$, entonces $Var(X) = (b-a)^2/12$.
- Si $X \sim N(\mu,\sigma^2)$, entonces $Var(X) = \sigma^2$.
- Si $X \sim$ Exp$(\lambda)$, entonces $Var(X) = 1/\lambda^2$.

- Hemos visto que, en todas las distribuciones estudiadas, los parámetros de las mismas están relacionados con sus *esperanzas* y sus *varianzas.*
- Un área importante de la estadística se encarga del diseño de métodos de estimación de los parámetros de las distribuciones.
- La vinculación entre *media, varianza* y los parámetros puede aprovecharse exitosamente en este sentido.

**Ejemplo**

Supongamos que para modelar una situación concreta se sabe que el modelo adecuado es una v.a. con distribución exponencial, pero se desconoce el parámetro $\lambda$. Una alternativa es estimarlo. Sabemos que si tomamos una muestra de la v.a. $X$, la media muestra $\bar{x}$ es 'parecida' a la media poblacional $E(X)$. Como $E(X) = 1/\lambda$, parece razonable estimar $\lambda$ con $1/\bar{x}$.

- Ya vimos que $\bar{x}$ (media muestral) estima $E(X)$ (media poblacional). Análogamente, la varianza muestral definida por

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}$$

es un buen estimador puntual de la varianza poblacional $Var(X)$ siempre que la muestra $x_1, \ldots, x_N$ sea lo suficientemente grande.