Theory of Probability   Oct 12, 2020

Expectation of X

$$E[X] = \sum_i x_i\, p(x_i) \qquad \text{when} \quad P[X = x_i] = p(x_i).$$

Consider this example:

Roll 1, get \$1

Roll 2, get \$4

$Y = X^2$, when X = roll of a die.

$$P[Y = 1] = P[X = 1] = 1/6$$

$$P[Y = 4] = P[X = 2] = 1/6$$

$$\vdots$$

$$E[Y] = \sum_i y_i\, p_Y(y_i) \longleftarrow$$

$$= \underbrace{\sum_i x_i^2\, p_X(x_i)}_{\text{expected value of } Y = X^2.}$$

In general, one can show that if $Y = g(X)$, then

$$\boxed{E[Y] = E[g(X)] = \sum_i g(x_i)\, p(x_i).}$$

Note: $E[X^2] \neq \left(E[X]\right)^2$

$$\left(E[X]\right)^2 = \left(\sum_i x_i\, p(x_i)\right)^2 = \sum_i \sum_j x_i x_j\, p(x_i)\, p(x_j) \neq \sum_i x_i^2\, p(x_i)$$

Corollary: Expectation is a linear operator:

$$E[aX+b] = \sum_i (ax_i+b)\,p(x_i)$$

$$= \underbrace{a\sum_i x_i\,p(x_i)}_{E[X]} + b\underbrace{\sum_i p(x_i)}_{1}$$

$$= aE[X] + b \quad \Rightarrow \quad E \text{ is a linear transformation.}$$

Remark: $E[X]$ is also known as the 1st moment of X.

Furthermore $E[X^n]$ is known as the nth moment of X:

$$E[X^n] = \sum_i x_i^n\,p(x_i).$$

Variance

Its useful to talk about characteristics of random variables:

- Expected value $\Leftrightarrow$ mean
- min value, max value
- "most likely value" $\rightarrow$ mode
- "spread" of a random variable.

Consider the following three R.V.s:

$P[W=0]=1$	$P[Y=-1]=\frac{1}{2}$	$P[Z=-100]=\frac{1}{2}$
$E[W]=0$	$P[Y=1]=\frac{1}{2}$	$P[Z=100]=\frac{1}{2}$
	$E[Y]=0$	$E[Z]=0$.

Let $\mu = E[X]$.

Want to characterize $E[|X-\mu|]$, but it turns out to be better mathematically to examine $E[(X-\mu)^2]$.

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

$$E[(X-\mu)^2] = \sum (x_i-\mu)^2 p(x_i)$$
$$= \sum (x_i^2 - 2\mu x_i + \mu^2) p(x_i)$$
$$= \underbrace{\sum x_i^2 p(x_i)}_{E[X^2]} - 2\mu \underbrace{\sum x_i p(x_i)}_{E[X]=\mu} + \mu^2 \underbrace{\sum p(x_i)}_{1}.$$
$$= E[X^2] - 2\mu^2 + \mu^2$$
$$= E[X^2] - \mu^2$$
$$= \underbrace{E[X^2]}_{2^{nd}\text{ moment}} - (\underbrace{E[X]}_{1^{st}\text{ moment}})^2$$

Variance is <u>not</u> a linear transformation.

$$\mathrm{Var}[aX+b] = E[(aX+b)^2] - (E[aX+b])^2$$
$$= E[a^2X^2 + 2abX + b^2] - (a\mu+b)^2$$
$$= a^2 E[X^2] + \cancel{2ab\mu} + \cancel{b^2} - a^2\mu^2 - \cancel{2ab\mu} - \cancel{b^2}$$
$$= a^2 (E[X^2] - (E[X])^2) = a^2 \mathrm{Var}[X]$$

Another useful (common) quantity is the <u>standard deviation</u>:

$$\mathrm{Std}[X] = \sqrt{\mathrm{Var}[X]}$$

$$\begin{aligned}\mathrm{Std}[aX+b] &= \sqrt{\mathrm{Var}[aX+b]} \\ &= \sqrt{a^2\,\mathrm{Var}[X]} \\ &= a\,\mathrm{Std}[X].\end{aligned}$$