Recall:- X is a discrete random variable

$X : S \to T$ $\quad E[X] = \sum_{t \in T} t \, P(X=t)$

Expectation of X ; mean of X ; Expected value of X ; Average value of X

$X = 10$ w.p. 1

$E[X] = 10$

$Y = \begin{cases} 9 & \text{wp } \frac{1}{2} \\ 11 & \text{wp } \frac{1}{2} \end{cases}$

$E[Y] = 9 \cdot \frac{1}{2} + 11 \cdot \frac{1}{2} = 10$

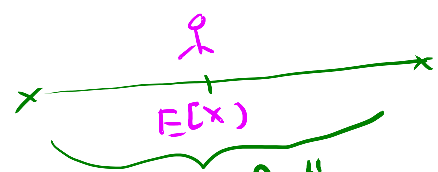
$Z = \begin{cases} 0 & \text{wp } \frac{1}{2} \\ 20 & \text{wp } \frac{1}{2} \end{cases}$

$E[Z] = 0 \cdot \frac{1}{2} + 20 \cdot \frac{1}{2} = 10$

$E(X) = E[Y] = E[Z] = 10.$

- Quantify how far away a random variable typically is from its average?

mean of X — notion of center



understand "spread"

Definition 4.2.1 Let X be a random variable with finite Expected value. Then the variance of the random variable X

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

The standard deviation of X is defined as

$$SD(X) := \sqrt{Var[X]}$$

[$Y \geq 0 \Rightarrow E(Y) \geq 0$: $SD[X]$ is well defined]

Observations:

$X : S \to T$

① $\mathrm{Var}[X] = \sum_{t \in T} (t - E[X])^2 \, \mathbb{P}(X = t)$

[ie $E[f(X)] = \sum_{t \in T} f(t) \, \mathbb{P}(X = t)$]

average of square distance of X from its Expected value.

If X has high probability of being far away from E[X] then the variance is large.

In our three cases X, Y, Z above:

$\mathrm{Var}[X] = 0$, $\mathrm{var}[Y] = \frac{(9-10)^2}{2} + \frac{(11-10)^2}{2} = 1$

$\mathrm{var}(Z) = \frac{(0-10)}{2} + \frac{(20-10)}{2} = 100$

$SD[X] = 0$, $SD[Y] = 1$

$SD[Z] = 10$

② X has associated units say meters. Then $\mathrm{Var}[X]$ has $(\text{meters})^2$ as its units

but $SD[X]$ has meters as its unit

$\therefore$ If $\mu := E[X]$ and $\sigma := SD[X]$ then

formally one can think of the Range(X)

$=$ "$(\mu - \sigma, \mu + \sigma)$"

③ $\mathrm{Var}[X] = \begin{cases} a & a \in \mathbb{R} \\ \infty & \end{cases}$ need not be finite

Understanding Spread of X from $\mu = E[X]$ and $\sigma = SD[X]$

$$k \in \mathbb{N} \qquad \mathbb{P}(|X-\mu| \geq k\sigma) = ?$$

- Given Pmf of X one can compute this explicitly.
- otherwise: Can we say how large or small this can be?

Theorem 4.14 (Chebyshev's inequality) X is a discrete random variable $\mu = E[X]$ and $\sigma = SD[X]$, $\mu < \infty$ and $\sigma < \infty$.. let $k > 0$

$$\mathbb{P}(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Proof:-

$$\{|X-\mu| \geq k\sigma\} = \{(X-\mu)^2 \geq k^2\sigma^2\}$$

let $Y = (X-\mu)^2$ $\quad Y \geq 0$ random variable (Discrete)

$c = k^2\sigma^2 \quad c > 0$

[Markov Inequality] $c > 0$ & $E(Y) < \infty$

$$\mathbb{P}(Y > c) \leq \frac{E[Y]}{c} \quad — \; (*)$$

Here: $E(Y) = E(X-\mu)^2 = Var[X] = \sigma^2 < \infty$

$c = k^2\sigma^2 > 0$

$(*) \Rightarrow \quad \mathbb{P}\left((X-\mu)^2 > k^2\sigma^2 \right) \le \frac{\sigma^2}{k^2\sigma^2}$

$\therefore \quad \mathbb{P}\left(|X-\mu| > k\sigma \right) \le \frac{1}{k^2} \qquad \forall\, k>0$

$\square$

observation
$k=3$

$$\mathbb{P}\left(X \notin (\mu-3\sigma, \mu+3\sigma) \right) \le \frac{1}{9}$$

Standardisation : X is a discrete random
$\mu = E[X]$ and $\sigma = SD[X]$

$0<\sigma<\infty$
$\mu<\infty$

Consider : $Z = \frac{X-\mu}{\sigma}$ is a discrete random variable

- $E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = E\left[\frac{X}{\sigma} - \frac{\mu}{\sigma}\right]$

$$= \frac{E[X]}{\sigma} - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

- $\mathrm{Var}[Z] = E\left[(Z-E[Z])^2\right]$

$$= E[Z^2] = E\left[\frac{(X-\mu)^2}{\sigma}\right]$$

$$= \frac{1}{\sigma^2} E[(X-\mu)^2] = \frac{\sigma^2}{\sigma^2} = 1$$

Theorem 4.2.4 : Let $a \in \mathbb{R}$, X be a random variable then $\mu = E[X] < \infty$ and $\sigma^2 = Var[X] < \infty$

(i) $Var[aX] = a^2 \, Var[X]$

(ii) $SD[aX] = a \, SD[X]$

(iii) $Var[X+a] = Var[X]$

(iv) $SD[X+a] = SD[X]$

Proof :- Ex D

Theorem 4.2.5 :- Let X be a random variable for which $E[X]$ and $E[X^2]$ are both finite, then

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

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

Now, $[X - E[X]]^2 = X^2 + (E[X])^2 - 2X\,E[X]$

$\therefore Var[X] = E[X^2 + (E[X])^2 - 2X\,E[X]]$

$= E[X^2] + (E[X])^2 - 2\,E[X\,E[X]]$

$\left(E[X]\right)^2$ is a constant) $= E[X^2] + \left(E[X]\right)^2 - 2E[X]\,E[X]$

$= E[X^2] - \left(E[X]\right)^2$

$\square$

Suppose X, Y are two discrete random variables

$X : S \to T$; $Y : S \to V$

$E[X] < \infty$; $E[Y] < \infty$

$SD[X] < \infty$; $SD[Y] < \infty$

Let $Z = X + Y$, shown earlier $E[Z] = E[X] + E[Y] < \infty$

$$Var[Z] = E\left[\left(Z - E[Z]\right)^2\right]$$

$$\overset{\text{Thm 4.2.5}}{=} E[Z^2] - \left(E[Z]\right)^2$$

Now,

$$Z^2 = (X+Y)^2 = X^2 + Y^2 + 2(XY)$$

[linearity of Expectation]

$$E[Z^2] = E[X^2] + E[Y^2] + 2E[XY]$$

$$\left(E[Z]\right)^2 = \left(E[X] + E[Y]\right)^2$$

$$= \left(E[X]\right)^2 + \left(E[Y]\right)^2 + 2E[X]\,E[Y]$$

$$\mathrm{Var}[Z] = E[Z^2] - (E[Z])^2$$
$$= E[X^2] + E[Y^2] + 2E[XY] - (E[X])^2 - (E[Y])^2 - 2E[X]E[Y]$$

$$\boxed{\begin{aligned} &\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y] \\ &\text{where } \mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y] \end{aligned}} \quad (\dagger)$$

Theorem 4.2.6: Suppose X and Y are INDEPENDENT random variables. (Assume $\mu_X, \mu_Y, \sigma_X, \sigma_Y < \infty$)

$$\mathrm{Var}(X+Y) = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

Proof: From $(\dagger)$ we know

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y]$$

But we have seen Thm 4.1.10: X, Y independent $E[XY] = E[X]E[Y]$ $\Rightarrow \mathrm{Cov}[X,Y] = 0$

X, Y independent $\Rightarrow \mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ $\square$

$f(x) = (x-p)^2$
$E(X-p)^2 = \sum_{t \in T} f(t) P(X=t)$

Example 4.2.8 $\quad X \sim \text{Bernoulli}(p)$

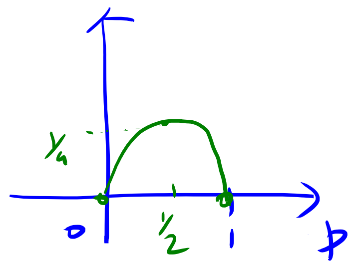
$0 \le p \le 1$

$$P(X=0) = 1 - P(X=1)$$
$$P(X=1) = p$$

$\therefore \; E[X] = p$

$$\text{Var}[X] = E[(X-p)^2]$$
$$= (0-p)^2(1-p) + (1-p)^2 p$$

$\text{Var}[X] = p(1-p)$



$$\therefore \; \text{Var}[X] = p^2(1-p) + (1-p)^2 p$$
$$= p^2 - p^3 + (1+p^2-2p)p$$
$$= p^2 + p - 2p^2 = p(1-p)$$

Example 4.2.9 $\quad X \sim \text{Binomial}(n, p) \qquad E[X] = np$

$$\text{Var}[X] = E(X-np)^2 = \sum_{k=0}^{n} (k-np)^2 \binom{n}{k} p^k (1-p)^{n-k}$$
$$= \cdots$$
$$= \cdots$$

$X = \sum_{i=1}^{n} \xi_i \qquad \xi_i = \text{independent \& identically distributed}$

i.i.d

$$\xi = \begin{cases} 0 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p \end{cases}$$

$\xi \sim \text{Bernoulli}(p)$

$\therefore$ Theorem 4.2.4 $\Rightarrow$

$$\text{Var}[X] = \sum_{i=1}^{n} \text{Var}(\xi_i)$$
$$= \sum_{i=1}^{n} p(1-p) = n p(1-p)$$

- Applications: • Sample from a population where some members have a certain characteristic and others do not.

Explore Soon.

Goal: to provide an estimate for the # of people with the characteristic.

- variance comes into play.

Remarks: let X be a continuous random variable with p.d.f. $f_X : \mathbb{R} \to [0, \infty)$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

- **Similarly:** under assumption that $E[X] < \infty$

$$Var[X] := E(X - E(X))^2 = \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x)\, dx$$

$$SD[X] = \sqrt{Var[X]}$$

- other properties also follow :-

$a > 0$ - $Var[aX] = a^2 Var[X]$

$Var[a+X] = Var[X]$

$E[X+Y] = E[X] + E[Y]$ (Probability II)

X, Y are independent: $Var[X+Y] = Var[X] + Var[Y]$

(Proof - next semester in Probability II)