

Recall.

Thm 4.119.

$X: S \rightarrow T$ discrete random variable

$$f: T \rightarrow U$$

$$u \in \mathbb{R}$$

$Y = f(X)$ is also a discrete random variable

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

Proof: "rearrangement of terms in a sum."

Thm 4.1.20 $X_1, X_2, \dots, X_n$ be a collection of discrete random variables on a common sample space S

$$X_j: S \rightarrow T_j \quad T_j \subseteq \mathbb{R} \text{ countable}, j=1, \dots, n$$

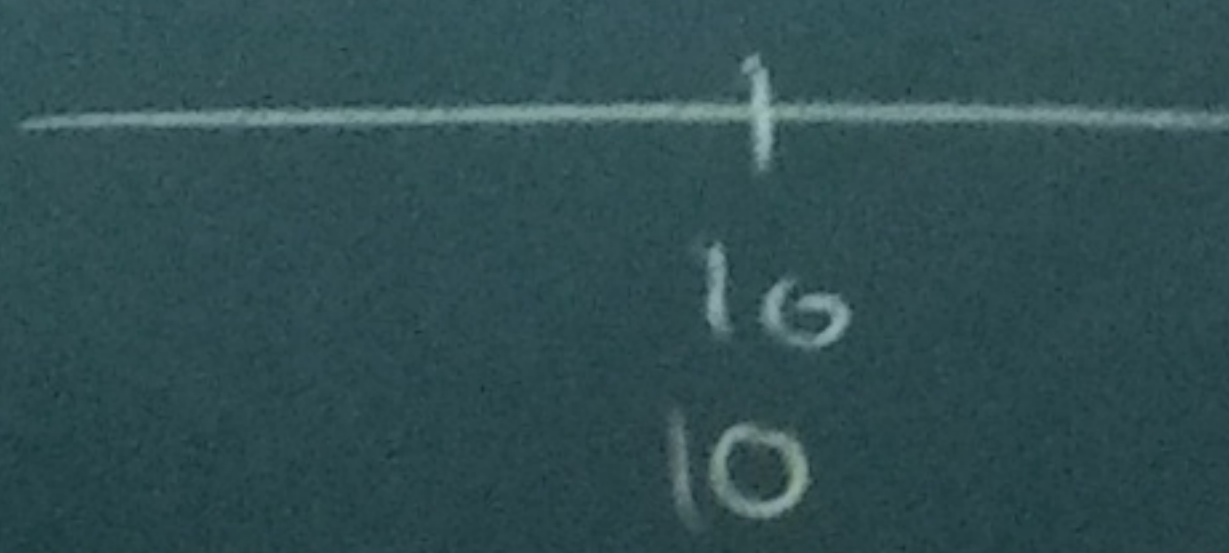
$$f: T_1 \times T_2 \times \dots \times T_n \rightarrow \mathbb{R}$$

$$E[f(X_1, \dots, X_n)] = \sum_{t_j \in T_j} f(t_1, \dots, t_n) P(X_1=t_1, X_2=t_2, \dots, X_n=t_n)$$

Proof. similar to proof of Theorem 4.1.19. $\square$

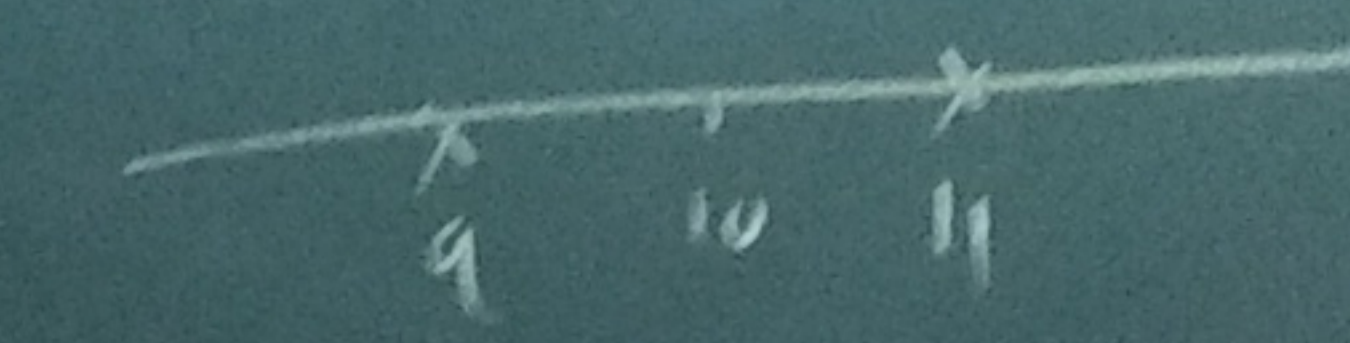
Standard deviation & Variance

X, Y, Z are three discrete random variables

$$X: S \rightarrow \mathbb{R}$$


$$X(\omega) = 10 \quad \forall \omega \in S$$

$$E[X] = 10P(X=10) = 10 \times 1 = \underline{\underline{10}}$$

$$Y: S \rightarrow \mathbb{R}$$


$$P(Y=9) = \frac{1}{2}$$

$$P(Y=11) = \frac{1}{2}$$

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

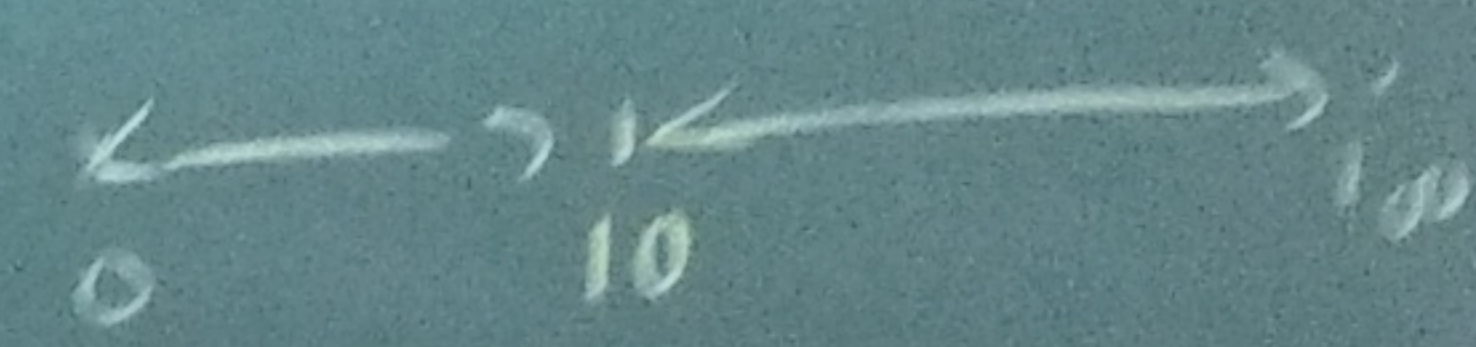
Useful to understand how the random variable
"spreads" around its expected value

$$Z: S \rightarrow \mathbb{R}$$

$$P(Z=0) = \frac{9}{10}$$

$$P(Z=100) = \frac{1}{10}$$

$$E[Z] = \underline{\underline{10}}$$



$X_n = t_n$

- average or expected value may not capture
key features of the random variable.

Definition 4.2.1 Let X be a discrete random variable with finite expected value. The VARIANCE of X , written as $\text{Var}[X]$, is defined as

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

(ie $f: T \rightarrow \mathbb{R}$ $f(t) = (t - E[X])^2$, $\text{Var}[X] = E[f(X)]$)

The STANDARD DEVIATION of a random variable, written as $\text{SD}[X]$, is defined as

$$\text{SD}[X] = \sqrt{\text{Var}[X]} \quad (\text{well defined})$$

$$\text{Var}[X] = E[X - E[X]]^2 = [10 - 10]^2 P(X=10) = 0 \quad \text{SD}(X) = 0$$

$$\text{Var}[Y] = E[Y - E[Y]]^2 = E[Y - 10]^2 = (9-10)^2 P(Y=9) + (11-10)^2 P(Y=11) = 1, \text{SD}[Y] = 1$$

$$E[f(Y)] = \sum_{y \in T} f(y) P(Y=y)$$

$$\begin{aligned} \text{Var}[Z] &= E[Z - E[Z]]^2 = (0-10)^2 P(Z=0) + (100-10)^2 P(Z=100) \\ &= 100 \cdot 9/10 + 8100 \cdot 1/10 = \frac{9000}{10} = 900 \quad \text{SD}[Z] = 30 \end{aligned}$$

- larger the variance the higher the probability that the random variable assumes values far from the expected value.

- Associated Units: for example if X is in meters, $\text{Var}[X] \equiv (\text{meters})^2$

- $\text{Var}[X]$ is a good computational tool. $\text{SD}[X]$ is in meters

- $(E[X] - \text{SD}[X], E[X] + \text{SD}[X])$ - in many cases there is an effective description of the Range(X).

- $E[X] \in \mathbb{R}$ but if $\text{SD}[X] = \infty$ means that any number under-estimate the spread of X from its Expected value.

Example 4.2.3

$$X \sim \text{Uniform}\{1, 2, 3, 4, 5, 6\}$$

$$E[X] = \sum_{i=1}^6 i P(X=i) = \frac{1}{6} \frac{6 \cdot 7}{2} = \frac{7}{2} = 3.5$$

$$\begin{aligned} \text{Var}[X] &= E[(X - E[X])^2] = \sum_{i=1}^6 (i - \frac{7}{2})^2 P(X=i) \\ &= \frac{1}{6} [25/4 + 9/4 + 1/4 + 1/4 + 9/4 + 25/4] = 35/12 \end{aligned}$$

$$SD(X) = \sqrt{\frac{35}{12}} \approx 1.7$$

- closest X can be to $E(X)$ is when it is 3 or 4
for each of these " $(X - E(X))$ " = 0.5

- farthest X can be to $E(X)$ is when it is 1 or 6
for each of these " $(X - E(X))$ " = 2.5

$$0.5 \leq SD(X) \leq 2.5$$

□

Theorem 4.2.4

$$a \in \mathbb{R}$$

X is a discrete random variable with finite

mean & variance.

$$(i) \text{Var}[aX] = a^2 \text{Var}[X] \quad , (iii) \text{SD}[aX] = |a| \text{SD}[X]$$

$$(ii) \text{Var}[X+a] = \text{Var}[X] \quad , (iv) \text{SD}[X+a] = \text{SD}[X]$$

Proof:

Thm 4.1.7 (a) $\Rightarrow E[aX] = aE[X]$, (b) $\Rightarrow E[X+a] = E[X] + a$

$$(i) \text{Var}[aX] = E[(aX - E[aX])^2] = E[a(X - E[X])^2] = a^2 E[(X - E[X])^2] = a^2 \text{Var}[X]$$

$$(ii) \text{Var}[X+a] = E[(X+a - E[X+a])^2] = E[(X - E[X])^2] = \text{Var}[X]$$