# Discrete Random Variables

**Notation**

- $\forall$ = for all
- $\sigma$ = standard deviation
- $E(X)$ = expected value
- Var = variance ($= \sigma^2$)

**Introduction**

- It is a random variable that will eventually take on a discrete value.
- Discrete means 'countable' whereas continuous is with an interval.

$$P(X = x)$$

the probability that a discrete random variable $X$ takes on some value, $x$

- A probability distribution lists all the values of $x$ & the probability of each occuring under $X$.

| $x$ | 0 | 1 | 2 | 3 | 4 | ... |
|---|---|---|---|---|---|---|
| $P(X = x)$ | 0.3 | 0.4 | 0.1 | 0.2 | 0 | ... |

- The rules for this are...
  - $\sum_{\forall x} p(x) = 1$ ($p(x)$ is the probability function)
  - $0 \leq p(x) \leq 1 \quad \forall x$

**Expected Value**

- It is the theoretical mean of the variable $\Rightarrow$ not based on sample data.

4, 5, 6

$E(X) = \underbrace{\mu}_{\text{mean}} = 5$

$\Rightarrow$ most of the time may not be a value that $X$ can take on

$$E(X) = \sum_{\forall x} \left( x \cdot p(x) \right) = \sum_{\forall x} \left( x \cdot P(X = x) \right)$$

(multiply each value by its probability of occurring & sum)

Also $E[g(x)] = \sum_{\forall x} g(x) \cdot p(x)$

Imagine simulating the discrete random variable $X$ millions of times, you would see the mean of all values converge towards $E(X)$

- $E(aX + b) = aE(X) + b$
- $E(X \pm Y) = E(X) \pm E(Y)$ ∵ the data size increase/decreases (only if they are independent) (proved on last page)
- $\bar{x}$ is the mean based on sample data

Variance $\sigma = \sqrt{\dfrac{[6 - 5]^2 + [5 - 5]^2 + [4 - 5]^2}{3}}$ mean on a sample of data ~ each squared to avoid negatives

$\sigma^2$ = variance ; above $\sigma$, most of the data is contained within one standard deviation of the mean, either side

- $Var = \sigma^2 = \dfrac{\sum (x - \mu)^2}{n} = E[(x-\mu)^2]$ ∵ $E(x) = \dfrac{\sum x}{n}$
- $\sigma^2 / Var$ is a measure of the variability in $X$.
- $\sigma^2 = E(X^2) - (E(X))^2$

"the mean of the squares minus the square of the means"

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

(Do not need to know this derivation)

$Var(aX+b) = a^2 Var(X)$ as by adding on a constant, it doesn't affect the variability of X.

$Var(X \pm Y) = Var(X) + Var(Y)$ (proved below)
(only if they are independent)

$$\Rightarrow Var(2X+3Y) = 2^2 Var(X) + 3^2 Var(Y)$$
$$= 4 Var(X) + 9 Var(Y)$$

Example for Variance

Let X = weight of cereal in cereal box & Y = weight of cereal in bowl

where $E(X) = 16\text{ oz}$, $\sigma_x = 0.8\text{ oz}$, $E(Y) = 4\text{ oz}$, $\sigma_y = 0.6\text{ oz}$
where $15 \leq X \leq 17$ & $3 \leq Y \leq 5$

$\Rightarrow 18 \leq X + Y \leq 22$, max value = heaviest box & heaviest bowl
$\therefore E(X+Y) = E(X) + E(Y) = 20\text{oz}$, the difference from the mean & the upper & lower bounds $\uparrow$ $\Rightarrow$ variance $\uparrow$ (2 either side compared to 1)

OR $10 \leq X - Y \leq 14$ where X is low, Y is high for min & vice versa for max.
SAME DIFFERENCE
$(15 - 5 = 10,\ 17 - 3 = 14)$

Like pythagoras in a way $\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2$

Example for $E(X)$

Let $X$ = # of dogs & $Y$ = # of cats

$$E(X) = \mu_x = 3 \quad \& \quad E(Y) = \mu_y = 4$$
$$\Rightarrow E(X+Y) = \mu_{x+y} = \mu_x + \mu_y = 7$$

- Imagine saying the expected no. of cats & dogs you'd see in a day

$$\Rightarrow E(Y-X) = \mu_{y-x} = \mu_y - \mu_x = 1$$

- Imagine the expected no. of more cats than dogs you'd see in a given day; expect to see 1 more cat than dogs.

Discrete Uniform Distribution

| $x$ | 1 | 2 | 3 | ... | $n$ |
|---|---|---|---|---|---|
| $P(X=x)$ | $1/n$ | $1/n$ | $1/n$ | | $1/n$ |

OR some transformation like:

| | | | | |
|---|---|---|---|---|
| 2 | 4 | 6 | ... | $n$ |
| $1/n$ | $1/n$ | $1/n$ | | $1/n$ |

$$E(T) = \frac{n+1}{2} \quad \& \quad Var(T) = \frac{n^2-1}{2}$$

Proof of $E(T)$ (need to know)

$$E(T) = (1 \times 1/n) + (2 \times 1/n) + (3 \times 1/n) + \dots + (n \times 1/n) \quad \because E(T) = \sum x P(X=x)$$
$$= 1/n\,(1 + 2 + 3 + \dots n)$$
$$= 1/n \left(\sum_{k=1}^{n} k\right) = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{\cancel{n}(n+1)}{2\cancel{n}} = \frac{n+1}{2}$$

Proof of $V(T)$

(need to know)

$$V(T) = E(T^2) - (E(T))^2$$

$$= \frac{1}{n} \times \sum_{k=1}^{n} k^2 - \left( \frac{1}{n} \times \sum_{k=1}^{n} k \right)^2$$ $\because$ probability $\times$ each term

$$= \frac{1}{\cancel{n}} \times \frac{\cancel{n}(n+1)(2n+1)}{6} - \left( \frac{1}{\cancel{n}} \times \frac{\cancel{n}(n+1)}{2} \right)^2$$

$$= \frac{(n+1)(2n+1)}{6} - \left( \frac{n+1}{2} \right)^2 = (n+1)\left[ \frac{2n+1}{6} - \frac{n+1}{4} \right]$$

$$= (n+1)\left[ \frac{4n+2}{12} - \frac{3n+3}{12} \right] = (n+1)\left[ \frac{n-1}{12} \right]$$

$$= \frac{(n+1)(n-1)}{12}$$

$$= \frac{n^2-1}{12}$$