On the importance of variance: Consider two RVs X and Y. Each takes only two values with equal probability.

X takes 495 and 505

$E(X) = \frac{1}{2} \cdot 495 + \frac{1}{2} 505 = 500$

$V(X) = \frac{1}{2}(495-500)^2 + \frac{1}{2}(505-500)^2$

$= \frac{1}{2} \cdot 5^2 + \frac{1}{2} \cdot 5^2 = 5^2 = 25$

$\sigma_X = \sqrt{25} = 5$

$c_X = \frac{5}{500} = 0.01$

Y takes 1 and 999

$E(Y) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 999 = 500$

$V(Y) = \frac{1}{2}(1-500)^2 + \frac{1}{2}(999-500)^2$

$= \frac{1}{2} \cdot 499^2 + \frac{1}{2} \cdot 499^2 = 499^2$

$\sigma_Y = 499$

$c_Y = \frac{499}{500} = 0.992$

Example:

Two players play a game by flipping a fair coin one or two times. If the first flip results in a head, then the game stops. Otherwise, the coin is flipped just one more time, and the game stops regardless of the outcome. If the game ends in a head, then player 1 loses \$1 to player 2. Otherwise, player 1 wins \$3 from player 2.

Give the sample space. If X is a RV denoting the winnings by player 1, give the values X takes.

Is it a fair game?

$\Omega = \{H, TH, TT\}$

with probabilities $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{4}$

$X(H) = -1$

$X(TH) = -1$

$X(TT) = 3$

$P(X=-1) = P(H \text{ or } TH \text{ occurs})$

$= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$

$P(X=3) = \frac{1}{4}$

$E(X) = (-1) \cdot p(-1) + 3\, p(3) = (-1) \cdot \frac{3}{4} + 3 \cdot \frac{1}{4} = 0$

Yes, it is a fair game.

Multiple RVs

Suppose X and Y are RVs defined on Ω. P(X=x, Y=y) is the joint probability mass function of the RVs.

Suppose X takes values $x_1, \ldots, x_m$, and Y takes values $y_1, \ldots, y_n$ with nonzero probabilities.

Their joint pmf can be given by a two-dimesional table.

X \ Y=	y_1	y_2	- - - i	y_n
x_1	$p(x_1, y_1)$			
x_2			$p(x_2, y_i)$	
⋮				
x_m				

Example: Consider a RE of rolling two 6-faced dice.

X : RV denoting the outcome of the first die

Y : " second die

$1 \le X \le 6, \quad 1 \le Y \le 6 \qquad |\Omega| = 36$

$$P(X=2, Y=3) = \frac{1}{36}$$

X \ Y	1	2	3	4	5	6	
1	1/36	1/36	1/36	1/36	1/36	1/36	$\Sigma = \frac{1}{6}\ P(X=1)$
2	1/36						$\Sigma = \frac{1}{6}\ P(X=2)$
3	1/36						
4	1/36						
5	1/36					1/36	
6	1/36						

$\Sigma\ 1/6 = P(Y=1)$

Independent Random Variables

If $P(X=x_i, Y=y_j) = P(X=x_i) * P(Y=y_j)$ for all values taken by X and Y, then

X and Y are independent RVs.

Functions of RVs

$Z = X + Y$ is a new RV on Ω

$W = X * Y$

$X_2 = 2 \cdot X$

$X_3 = X^2$

Properties of Expectation and Variance

Let X and Y be two RVs. Let c be any constant.

Expectation

(i) $E(c) = c$

(ii) $E(c \cdot X) = c \cdot E(X)$

(iii) $E(X+Y) = E(X) + E(Y)$ ← linearity property

(iv) $E(X \cdot Y) = E(X) \cdot E(Y)$ $\underbrace{\text{if } X \text{ and } Y \text{ are independent}}_{X \perp Y}$

Variance

(i) $V(c) = 0$

(ii) $V(cX) = c^2 V(X)$

(iii) $V(X+Y) = V(X) + V(Y)$ if X and Y are indep.

$V(X-Y) = V(X) + V(Y)$

$V(aX+bY) = a^2 V(X) + b^2 V(Y)$, $X \perp Y$

a, b are constants

$V(X) = E(X^2) - [E(x)]^2$

Some interesting discrete distributions

Bernoulli distribution

A RE that has exactly two outcomes is a Bernoulli trial.

Success with prob. p

Failure with prob. $1-p$

Examples: 1. A coin flip with probability of head (success) p.

2. A conditional statement: if (B) then {...} else {...}

Let X be a RV describing the outcome of a Bernoulli trial.

$\Rightarrow$ X is a Bernoulli RV

$\Rightarrow$ $X \sim \text{Bernoulli}(p)$

X takes two values: 0 (failure), 1 (success)

$$\boxed{\begin{array}{l} p(0) = P(X=0) = 1-p, \qquad p(1) = P(X=1) = p \\ p(x) = 0 \quad \text{if } x \neq 0 \text{ or } 1 \end{array}}$$

$\leftarrow$ pmf

Verify: $\sum_{x_i} p(x_i) = p(0) + p(1) = (1-p) + p = \underline{\underline{1}}$

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

$F(1) = P(X \le 1) = P(X \le 0) + P(X=1) = F(0) + p(1) = (1-p) + p = 1$

$E(X) = \sum_{x_i} x_i\, p(x_i) = 0 \cdot \underbrace{p(0)}_{(1-p)} + 1 \cdot \underbrace{p(1)}_{p} = p = \mu$

$\sigma^2 = V(X) = \sum_{x_i} (x_i - \mu)^2 p(x_i) = (0-p)^2 \cdot (1-p) + (1-p)^2 p = p(1-p)$

$C = \sigma/\mu = (q/p)^{1/2}$

Example: Consider a fair coin flip. P(H) = P(T) = 1/2 . Let X be the RV indicating the number of heads.

Then $X \sim \text{Bernoulli}(1/2)$. $\mu = E(X) = p = 1/2$ $\sigma^2 = V(X) = p(1-p) = 1/2 * 1/2 = 1/4$

$\sigma = (1/4)1/2 = 1/2$ $C = \sigma/\mu = 1$

Binomial distribution

Consider **n** independent Bernoulli trials each with probability of success **p**.
Let Y denote the number of successes in n trials. Then Y takes values 0, 1, 2, ..., n.

Y is a binomial RV.
Notation: Y ~ Binomial(n,p).

$$(x+y)^n = \binom{n}{0} x^0 y^n + \binom{n}{1} x^1 y^{n-1} + \cdots + \binom{n}{n} x^n y^0$$

$$P(Y=0) = p(0) = P(\text{failure in trial } 1) \cdots P(\text{fail in trial } n)$$
$$= (1-p)\cdot(1-p) \cdots (1-p) = (1-p)^n.$$

$$p(1) = P(Y=1) = \binom{n}{1} \cdot p(1-p)^{n-1}$$

$$\begin{array}{ccccc} p & (1-p) & (1-p) & \cdots & (1-p) \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 3 & & n \end{array}$$

$$p(1-p)^{n-1}$$

$$p(2) = P(Y=2) = \binom{n}{2} p^2 (1-p)^{n-2}$$

often denoted $b(k;n,p)$

pmf

$$p(k) = P(Y=k) = \begin{cases} \overbrace{\binom{n}{k} p^k (1-p)^{n-k}} & \text{for } 0 \le k \le n,\ k \text{ is an integer} \\ 0 & \text{otherwise} \end{cases}$$

pmf: $\sum_{k=0}^{n} p(k) = 1 \quad \leftarrow$ verify.

$$F(0) = p(0) = (1-p)^n = \binom{n}{0} p^0 (1-p)^n$$

$$F(1) = P(X \le 1) = P(X=0) + P(X=1) = \binom{n}{0} p^0 (1-p)^n + \binom{n}{1} p (1-p)^{n-1}$$

$$\underbrace{F(k)}_{\text{denoted } B(k;n,p)} = P(X \le k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$$

Y ~ Binomial(n,p). $\mu = E(Y) = np$ $\sigma^2 = V(Y) = npq$ $\sigma = (npq)^{1/2}$
q=1-p $C = \sigma/\mu = [q/(np)]^{1/2}$

Example: A sealed box has 47 Crackers. Each cracker may be broken with probability 0.3 independently of the other crackers in the box.

(i) What is the probability that exactly 18 crackers are broken?

$n = 47 \qquad p = 0.3$

X – RV for the number of broken crackers

$$P(X=18) = p(18) = \binom{47}{18} 0.3^{18}\, 0.7^{47-18}$$

(ii) What is the probability that 2 or more crackers are broken?

$$P(X=2) + P(X=3) + \cdots + P(X=47)$$

$$\sum_{k=2}^{47} \binom{n}{k} 0.3^{k}\, 0.7^{47-k} = 1 - F(1)$$

$$P(X \le 1) = F(1) = 0.7^{47} + \binom{47}{1} 0.3\; 0.7^{46}$$

$$\sum_{k=0}^{47} \binom{n}{k} 0.3^{k}\, 0.7^{47-k} = 1$$

(iii) What is the average, expected, number of broken Crackers in the box?

$$\mu = E(X) = \sum_{k=0}^{n} k \cdot p(k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = np$$

$$= 47 * 0.3 = 14.1$$

(iv) What is the standard deviation of the number of broken crackers in the box?

$$V(X) = \sum_{k=0}^{n} (k-\mu)^2 p(k) = \sum_{k=0}^{n} (k-np)^2 \binom{n}{k} p^k (1-p)^{n-k} = npq$$

$$= 47 * 0.3 * 0.7 = 9.87$$

$$\sigma = \sqrt{9.87} = 3.14$$

Geometric Distribution

Let Z be the RV. that counts the number of Bernoulli trials until the first success occurs.

If 0 is failure, 1 is success

$\Omega = \{ 0^i 1 \mid i = 0, 1, 2, \ldots \}$

$\Omega = \{ H, TH, TTH, \ldots \}$

1   01   001   -   -

Let p be the prob. of success in the Bernoulli trial.

pmf: $P(Z=i) = (1-p)^{i-1} \cdot p, \quad i = 1, 2, \ldots$

CDF: $F(i) = \sum_{j=1}^{i} (1-p)^{j-1} \cdot p, \quad i \geq 1$

$F(i) = 1 - (1-p)^i, \quad i \geq 1$

$P(Z > k) = (1-p)^k p + (1-p)^{k+1} p + \cdots = (1-p)^k \cdot p \cdot \left[\frac{1}{1-(1-p)} \right] = (1-p)^k$

Z ~ Geometric(p) $\mu = E(Z) = 1/p$ $\sigma^2 = V(Z) = q/p^2$ $\sigma = q^{1/2}/p$

q=1-p $C = \sigma/\mu = (q/p)^{1/2}$