Probability and Events

01 October 2025 19:06

- Consider an experiment with sample space $S = \{1, 2, \dots, m\}$.
  Suppose that there are $p_1, \dots, p_m$ with :

  $$p_i \geq 0 \ , \ i = 1, 2, \dots, m \quad \& \quad \sum_{i=1}^{m} p_i = 1$$

  p_i is probability that i is the out come.

- Example : Find probabilities of Head & Tail of a unbiased coin

  Sample space $S = \{H, T\}$ $\therefore$ $P(H) = P(T)$ (Unbiased coin)

  & $P(H) + P(T) = 1 \Rightarrow P(H) = P(T) = 1/2$

- Example : Two dice are rolled (fair). Prob that sum is 7.

  Sample space $S = \{(x, y) \mid x, y \in \{1, 2, \dots, 6\}\}$

  $\therefore$ For any $(x, y) \in S$ $P(x, y) = 1/36$

  Favorable outcomes $A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

  $$P(A) = P(1,6) + P(2,5) + \dots + P(6,1) = 6 \times 1/36 = \underline{\underline{1/6}}$$

- Note : $A^c :=$ complement of A

  (elements in S which are not in A)

  $$\sum_{i \in S} p_i = 1 \Rightarrow \sum_{i \in A} p_i + \sum_{i \in A^c} p_i = 1$$

$$\sum_{i \in S} \quad \sum_{i \in A} \quad \sum_{i \in A^c}$$

$$\Rightarrow \quad P(A) + P(A^c) = 1 \Rightarrow P(A^c) = 1 - P(A)$$

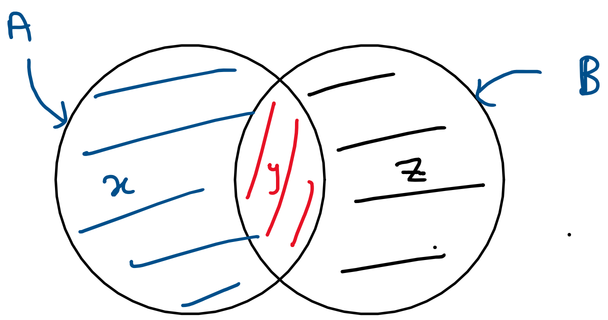
- $P(S) = 1, \quad S^c = \phi \Rightarrow P(\phi) = 0$

Proposition

Let $A, B \in \mathcal{P}(S)$ be event in S, then

$$P(A \cup B) + P(A \cap B) = P(A) + P(B)$$

Pf:



$P(A \cup B) = x + y + z$

$P(A \cap B) = y$

$P(A \cup B) + P(A \cap B) = x + 2y + z$ —— ①

$\left.\begin{array}{l} P(A) = x + y \\ P(B) = y + z \end{array}\right\} P(A) + P(B) = x + 2y + z$ —— ②

from ① & ②, LHS = RHS $\square$

- **Example**: P(Stock increase Today) = 0.54

P(Stock increase Tommo.) = 0.54

P(Stock incr both days) = 0.28

P(stock doesn't increase) = ?

A = Stock increase today

B = Stock increase tomorrow

$P(A \cup B) + P(A \cap B) = P(A) + P(A)$

$P(A \cup B) + 0.28 = 0.54 + 0.54 \Rightarrow P(A \cup B) = 0.80$

$\therefore P(A \cup B) =$ P(Stock increase on any day) $= 0.80$

P(stock doesn't increase) $= 1 - 0.80 = \underline{0.2}$

Note:

If A & B are independent i.e. $A \cap B = \phi$

then $P(A \cup B) = P(A) + P(B)$

Conditional Probability

01 October 2025 19:37

- Let $A, B \in \mathcal{P}(S)$ be events in S, then

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

- Example : Toss two coin :

$P(H, H) = 1/4$

$$P(\text{Second is Head} \mid \text{First was H}) = \frac{|\{(H, H)\}|}{|\{(H,H), (T, H)\}|} = \underline{\underline{1/2}}$$

Or

$$\frac{P(A \cap B)}{P(A)} = \frac{P(H, H)}{P((H,T), (H,H))} = \frac{1/4}{1/2} = \underline{\underline{1/2}}$$

Example 1.2a A coin is flipped twice. Assuming that all four points in the sample space $S = \{(h, h), (h, t), (t, h), (t, t)\}$ are equally likely, what is the conditional probability that both flips land on heads, given that

(a) the first flip lands on heads, and
(b) at least one of the flips lands on heads?

Soln :

(a) $A = \{(H,H), (H,T)\} \Rightarrow P(A) = \frac{2}{4} = \underline{\underline{1/2}}$

$B = \{(H,H)\}$

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(H,H)}{P(A)} = \frac{1/4}{1/2} = \underline{\underline{1/2}}$$

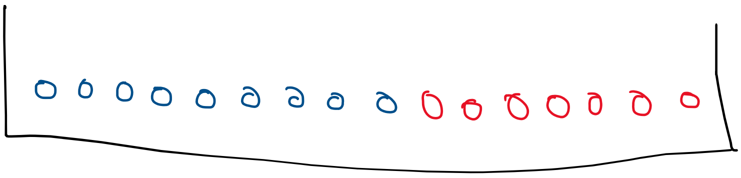
(b) $A = \{(H,H), (H,T), (T,H)\} \Rightarrow P(A) = 3/\cancel{4}$

(b) $A = \{(H,H),(H,T),(T,H)\} \Rightarrow P(A) = 3/\cancel{8}\,4$

$B = \{(H,H)\}$

$P(B|A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{P(H,H)}{P(A)} = \dfrac{1/4}{3/\cancel{8}\,4} = \underline{1/3}$

Example 1.2b Suppose that two balls are to be withdrawn, without replacement, from an urn that contains 9 blue and 7 yellow balls. If each ball drawn is equally likely to be any of the balls in the urn at the time, what is the probability that both balls are blue?



$P(\text{Both ball are blue}) = \dfrac{\cancel{9}\,3}{\cancel{16}\,2} \times \dfrac{\cancel{8}}{\cancel{15}\,5} = \underline{3/10}$

OR

B_1 = first ball is blue = $\{(B,B),(B,Y)\}$

B_2 = second ball is blue = $\{(B,B),(Y,B)\}$

$P(B_2|B_1) = \dfrac{P(B_1 \cap B_2)}{P(B_1)} \Rightarrow P(B_1 \cap B_2) = P(B_1) \times P(B_2|B_1)$

$= \dfrac{9}{16} \times \dfrac{8}{15} = \underline{3/10}$

Note

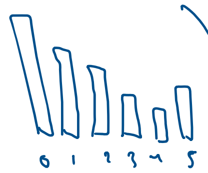
A, B are events in S, s.t. A & B are independent

then $P(A|B) = P(A) \Rightarrow P(A|B) = \dfrac{P(A \cap B)}{P(B)} = P(A)$

$\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

Example 1.2c Suppose that, with probability .52, the closing price of a stock is at least as high as the close on the previous day, and that the results for succesive days are independent. Find the probability that the closing price goes down in each of the next four days, but not on the following day.

Sol^n:

$P\Big($ 

$\Big) = P(0\to1) \cdot P(1\to2) \cdots P(4\to5)$

$= 0.48 \times 0.48 \times 0.48 \times 0.48 \times 0.52$

$= 0.028$

Random Variables and Expected Values

01 October 2025 20:09

- Numerical quantities whose values are determined by the outcome of a experiment are known as random variables.
- Since the value of a random variable is determined by a outcome of a experiment, we can assign probabilities to them.
- If X is a random variable whose possible values are $x_1, x_2, \cdots, x_n$ then the set $\{P(X = x_j) \mid j = 1, 2, \cdots, n\}$ is called probability distribution of X.
- $\sum_{i=1}^{n} P(X = x_i) = 1$

> Definition: If X is a r.v. whose possible values are $x_1, x_2, \cdots, x_n$, then the expected value of X, denoted by $E(X) := \sum_{j=1}^{n} x_j P(X = x_j)$.

Example 1.3b Let the random variable X denote the amount that we win when we make a certain bet. Find $E[X]$ if there is a 60% chance that we lose 1, a 20% chance that we win 1, and a 20% chance that we win 2.

Soln: $X = \begin{cases} -1 & 0.6 \\ 1 & 0.2 \\ 2 & 0.2 \end{cases}$

$E(X) = -0.6 + 0.2 + (0.2)2 = 0$

* Bernoulli Distb.

- Let X be r.v. such that $X = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$

then $X \sim \text{Bern}(p)$

- $E(X) = p(1) + (1-p)(0) = p$

> Note: X is a r.v and a, b are consts.
>
> $E(aX+b) = aE(X) + b$.
>
> Pf: $E(aX+b) = \sum_{i=1}^{n} (ax_i + b) P(X = x_i) = a\sum x_i P(X = x_i) + b \sum P(X = x_i)$
>
> $= a E(X) + b(1) = \underline{aE(X) + b}$

> Proposition
>
> $$E\left[\sum_{i=1}^{k} X_i\right] = \sum_{i=1}^{k} E(X_i)$$

* Binomial Random Variable

- Consider n independent trials, each of which is a success with prob p. A r.v. X = no. of success is called binomial r.v. with parameters n, p.
- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
- $E(X) = \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$
- $X = X_1 + X_2 + \cdots + X_n \quad X_i \sim \text{Bern}(p)$
- $E(X) = n E(X_i) = \underline{np}$

- $X = X_1 + X_2 + \cdots + X_n$
- $E(X) = n\,E(X_i) = np$

Proposition 1.3.2 *Consider n independent trials, each of which is a success with probability p. Then, given that there is a total of i successes in the n trials, each of the $\binom{n}{i}$ subsets of i trials is equally likely to be the set of trials that resulted in successes.*

Pf: $X \sim Bin(n,p)$

Let T be subset of i trials. A be event such that all trials in T were successfull.

Claim: $\dfrac{P(A \mid X=i)}{P(X=i)} = \dfrac{1}{\binom{n}{i}}$

$$\frac{P(A\mid X=i)}{P(X=i)} = \frac{\cancel{p^k (1-p)^{n-k}}}{\binom{n}{i}\cancel{p^k (1-p)^{n-k}}} = \frac{1}{\binom{n}{i}}$$

☆ Variance

- Definition: The variance of r.v. X, denoted by $Var(X)$ or $V(X)$ is defined by $Var(X) = E[(X - E(X))^2]$

Example 1.3f Find $\text{Var}(X)$ when X is a Bernoulli random variable with parameter p.

Sol^n: $X = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$ $\quad E(X) = p$

$$(X - E(X))^2 = \begin{cases} (1-p)^2 & p \\ p^2 & 1-p \end{cases}$$

$$E((X - E(X))^2) = p(1-p)^2 + (1-p)p^2 = (1-p)p[1-p+p]$$

$$Var(X) = (1-p)p$$

* Note: a, b are const. X is r.V., then

$$Var(aX+b) = a^2 Var(X).$$

Pf: $Var(aX+b) = E\left((aX+\cancel{b} - (aE(X)+\cancel{b}))^2\right)$

$= a^2 E((X - E(X))^2)$

$= a^2 Var(X)$

Proposition

Let $X_1, X_2, \ldots, X_n$ be indp^n. r.v.'s & $a_1, a_2, \ldots, a_n$ be const.

$$Var\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i^2 Var(X_i).$$

Example 1.3g Find the variance of X, a binomial random variable with parameters n and p.

Sol^n: $X \sim Bin(n,p)$ then let $X_i \sim Bern(p)$ s.t.

$$X = \sum_{i=1}^n X_i.$$

$$Var(X) = \sum_{i=1}^n Var(X_i) = np(1-p)$$

- Square root of variance is called standard deviation.

Covariance and Correlation

02 October 2025 02:50

- The covariance of any two r.v.'s X & Y is denoted by

$$Cov(X,Y) := E\left([X-E(X)][Y-E(Y)]\right)$$

$$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$= E(XY) - E(X)E(Y) - \cancel{E(X)E(Y)} + \cancel{E(X)E(Y)}$$

$$= E(XY) - E(X)E(Y)$$

- A positive cov( ) suggests that both r.v.'s tend to be large at same time.
- A negative cov( ) suggests that when one is large, the other tends to be small.
- If X & Y are independent $\Rightarrow$ $Cov(X,Y) = 0$.

* **Note**:
  - $Cov(X,Y) = Cov(Y,X)$
  - $Cov(X,X) = Var(X)$
  - $Cov(aX,Y) = a\,Cov(X,Y)$
  - $Cov(a,Y) = 0$
  - $Cov(X_1+X_2,Y) = Cov(X_1,Y) + Cov(X_2,Y)$

  Pf:

$$Cov(X_1+X_2,Y) = E(X_1Y + X_2Y) - E(X_1+X_2)E(Y)$$

$$= [E(X_1Y) - E(X_1)E(Y)] + [E(X_2Y) - E(X_2)E(Y)]$$

$$= Cov(X_1,Y) + Cov(X_2,Y)$$

$$\left[E(X_2 Y) - E(X_2)E(Y)\right]$$

$$= Cov(X_1, Y) + Cov(X_2, Y)$$

- $$Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \, Cov(X_i, Y_j)$$

- $$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{i \neq j} Cov(X_i, X_j)$$

Pf:

$$Var\left(\sum_{i=1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n}\sum_{j=i} Cov(X_i, X_j) + \sum_{i=1}^{n}\sum_{j \neq i} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j \neq i} Cov(X_i, X_j)$$

- The degree to which large values of X associate with large values of Y is measured by correlation of X & Y, denoted by

$$\rho(X, Y) := \frac{Cov(X, Y)}{\sqrt{Var(X)\, Var(Y)}}$$

- $$-1 \leq \rho(X, Y) \leq 1$$

Pf: $|\rho(X,Y)| \leq 1 \Rightarrow \rho^2(X,Y) \leq 1$ (To prove)

for a real t.

Assume $Y = tX$

$$\rho^2(X,Y) = \frac{Cov^2(Y, tX)}{(X)\, Var(Xt)} = \frac{t^2 \cancel{Var^2(X)}}{\cancel{t^2 Var^2(X)}} = 1 \leq 1$$

$$\rho^2(X,Y) = \frac{\text{Cov}(Y, tX)}{\text{Var}(X)\,\text{Var}(Xt)} = \frac{[illegible]}{\cancel{t^2\,\text{Var}^2(X)}} = [illegible]$$

Assume $Y \neq tX$

$$0 < \text{Var}(Y - tX) = \text{Var}(Y) - 2t\,\text{Cov}(X,Y) + t^2\,\text{Var}(X)$$

Discriminant for t should be negative.

$$4\,\text{Cov}^2(X,Y) - 4\,\text{Var}(Y)\,\text{Var}(X) < 0$$

$$\frac{\text{Cov}^2(X,Y)}{\text{Var}(X)\,\text{Var}(Y)} < 1 \Rightarrow \rho^2(X,Y) < 1 \qquad \blacksquare$$

- Let $Y = a + bX$

$$\rho(Y,X) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}} = \frac{\text{Cov}(X, a+bX)}{\sqrt{\text{Var}(X)\cdot\text{Var}(a+bX)}}$$

$$= \frac{b\,V(X)}{|b|\,V(X)} = \frac{b}{|b|}$$

$$\therefore \rho(X,Y) = \begin{cases} 1 & b > 0 \\ 0 & b = 0 \\ -1 & b < 0 \end{cases}$$

Conditional Expectation

02 October 2025 03:32

- for r.v.'s X & Y we define Condition expectation of X given $Y=y$ by

$$E(X|Y=y) = \sum_x x\, P(X=x|Y=y)$$

★ **Proposition:**

$$E(X) = \sum_y E(X|Y=y)\, P(Y=y)$$

Pf:

$$\sum_y E(X|Y=y)\, P(Y=y) = \sum_y \sum_x x\, P(X=x|Y=y)\, P(Y=y)$$

$$= \sum_y \sum_x x\, P(X=x \text{ and } Y=y)$$

$$= \sum_x x \sum_y P(Y=y \text{ and } X=x)$$

$$= \sum_x x P(X=x)$$

$$= E(X)$$

- $$E(h(Y)) = \sum_y h(y)\, P(Y=y)$$

- $$E(E(X|Y=y)) = \sum_y E(X|Y=y) \cdot P(Y=y)$$

  Put $h(y) = E(X|Y=y)$.

- $$E(X) = E(E(X|Y=y))$$

  Or simply, $E(X) = \underline{E(E(X|Y))}$

Exercise 1.1 When typing a report, a certain typist makes i errors with probability p_i $(i \geq 0)$, where

$$p_0 = .20, \quad p_1 = .35, \quad p_2 = .25, \quad p_3 = .15.$$

What is the probability that the typist makes

(a) at least four errors;
(b) at most two errors?

(a) P (At least fore errors) = 1 − P(less than four errors)
= 1 − (0.2 + 0.35 + 0.25 + 0.15)
= 1 − 0.95
= 6.05

(b) P (at most two errors) = P(0 error) + P(1 error) + P(2 error)
= .2 + .35 + .25
= 0.8

Exercise 1.2 A family picnic scheduled for tomorrow will be postponed if it is either cloudy or rainy. If the probability that it will be cloudy is .40, the probability that it will be rainy is .30, and the probability that it will be both rainy and cloudy is .20, what is the probabilty that the picnic will not be postponed?

Let A & B be events s.t.

A = rainy tomorrow
B = cloudy tomorrow

need to find $P((A \cup B)^c) = 1 - P(A \cup B)$
$= 1 - (P(A) + P(A) - P(A \cap B))$

$= 1 - 0.5 = 0.5$

Exercise 1.3 If two people are randomly chosen from a group of eight women and six men, what is the probability that

(a) both are women;
(b) both are men;
(c) one is a man and the other a woman?

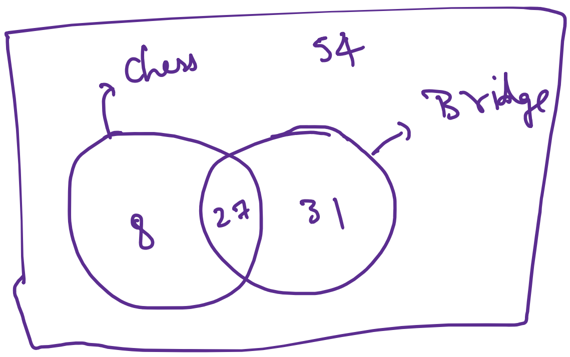
(a) $\frac{8\times 7}{^{14}C_2} = \frac{8\times \cancel{7}}{\cancel{2}\,\cancel{14}\times 13}\times \cancel{2} = \underline{\underline{8/13}}$

(b) $\frac{6\times 5}{^{14}C_2} = \frac{{}^{3}\cancel{6}\times 5}{7\,\cancel{14}\times 13}\times 2 = \underline{\underline{30/91}}$

(c) $\frac{6\times 8}{^{14}C_2} = \frac{{}^{3}\cancel{6}\times 8\times 2}{\cancel{2}\,\cancel{14}\times 13} = \underline{\underline{48/91}}$

Exercise 1.4 A club has 120 members, of whom 35 play chess, 58 play bridge, and 27 play both chess and bridge. If a member of the club is randomly chosen, what is the conditional probability that she

(a) plays chess given that she plays bridge;
(b) plays bridge given that she plays chess?



(a) $\frac{27}{58}\left[\frac{P(\text{Chess and Bridge})}{P(\text{Bridge})}\right]$

(b) $\frac{27}{35}\left[\frac{P(\text{Chess and Bridge})}{P(\text{Chess})}\right]$

Exercise 1.5 Cystic fibrosis (CF) is a genetically caused disease. A child that receives a CF gene from each of its parents will develop the disease either as a teenager or before, and will not live to adulthood. A child that receives either zero or one CF gene will not develop the disease. If an individual has a CF gene, then each of his or her children will independently receive that gene with probability 1/2.

(a) If both parents possess the CF gene, what is the probability that their child will develop cystic fibrosis?

Exercise 1.5 Cystic fibrosis (CF) is a genetically caused disease. A child that receives a CF gene from each of its parents will develop the disease either as a teenager or before, and will not live to adulthood. A child that receives either zero or one CF gene will not develop the disease. If an individual has a CF gene, then each of his or her children will independently receive that gene with probability 1/2.

(a) If both parents possess the CF gene, what is the probability that their child will develop cystic fibrosis?

(b) What is the probability that a 30-year old who does not have cystic fibrosis, but whose sibling died of that disease, possesses a CF gene?

(a) P(Develop CF) = P(Get CF gene from mother and get CF gene from father)

$= 1/2 \times 1/2 = \underline{\underline{1/4}}$

(b) Because his/her sibling died of CF, both mother & father had one gene each.

P(did not develop CF / Mother & Father has CF genes)

= P(has 0 CF gene) + P(has 1 CF gene from mother) + P (1 gene from father

$= 1/2 \times 1/2 + 1/2 \times 1/2 + 1/2 \times 1/2$

$= 3/4$

P(Posses a gene | Doesn't have disease)

$= \dfrac{P(\text{gene} = cC)}{P(\text{gene} = cC \text{ or } CC)} = \dfrac{1/2}{3/4} = \underline{\underline{2/3}}$

★ **Exercise 1.6** Two cards are randomly selected from a deck of 52 playing cards. What is the conditional probability they are both aces, given that they are of different suits?

P(both are aces | both cards from different suit)

$= \frac{4 \times 3}{\dots} = 2 \times 3 \times 4$

$P($ both are aces $|$ both ...

$P(A) = P(\text{both are aces}) = \dfrac{4\times 3}{{}^{52}C_2} = \dfrac{2\times 3\times 4}{52\times 51}$

$P(B) = P(\text{both are from diff suit}) = \dfrac{52\times 39\times 2}{52\times 51}$

$P(A\cap B) = P(A) \quad \because A \subset B$

$\Rightarrow P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{P(A)}{P(B)} = \dfrac{\cancel{2}\times\cancel{3}\times\cancel{4}}{\cancel{2}\times\cancel{39}_{13}\times\cancel{52}_{13}} = \dfrac{1}{169}$

Exercise 1.7 If A and B are independent, show that so are

(a) A and B^c;
(b) A^c and B^c.

given : A & B are indp' $\Leftrightarrow$ $P(A|B) = P(A)$ or $P(B|A) = P(B)$

(a) $P(B^c) = 1 - P(B) = 1 - P(B|A) = P(B^c|A)$ $\boxed{2}$

(b) from (a), $P(A) = P(A|B^c)$ $\boxed{2}$

$P(A^c) = 1 - P(A) = 1 - P(A|B^c) = P(A^c|B^c)$

OR

given: $P(A\cap B) = P(A)\cdot P(B)$

(a) $A = (A\cap B) \cup (A\cap B^c)$

$P(A) = P(A\cap B) + P(A\cap B^c)$

$P(A) = P(A).P(B) + P(A\cap B^c)$

$P(A)[1 - P(B)] = P(A\cap B^c)$

$P(A)P(B^c) = P(A\cap B^c)$ $\boxed{2}$

(b) $B^c = (B^c\cap A) \cup (B^c\cap A^c)$

$$P(B^c) = P(B^c \cap A) + P(B^c \cap A^c)$$

$$P(B^c) = P(B^c)\,P(A) + P(\,B^c \cap A^c)$$

$$P(B^c)\,P(A^c) = P(A^c \cap B^c) \quad \boxtimes$$

Exercise 1.8 A gambling book recommends the following strategy for the game of roulette. It recommends that the gambler bet 1 on red. If red appears (which has probability $18/38$ of occurring) then the gambler should take his profit of 1 and quit. If the gambler loses this bet, he should then make a second bet of size 2 and then quit. Let X denote the gambler's winnings.

(a) Find $P\{X > 0\}$.
(b) Find $E[X]$.

(a)
$$X = \begin{cases} 1 & 18/38 + \frac{20}{38} \times \frac{18}{38} = \frac{18 \times 58}{38^2} \\ -3 & \frac{20}{38} \times \frac{20}{38} = \frac{20 \times 20}{38^2} \end{cases}$$

$$P(X > 0) = P(X = 1) = \frac{18 \times 58}{38^2} = 0.723$$

(b)
$$E(X) = \frac{18 \times 58}{38^2} - 3 \cdot \frac{20 \times 20}{38^2} = \left(1 - \frac{20 \times 20}{38^2}\right) - 3\frac{}{}$$

$$= 1 - 4 \cdot \frac{20 \times 20}{(38)^2} = 1 - \frac{1600}{1444} = \underline{0.108}$$

$$= \underline{0.108}$$

Exercise 1.9 Four buses carrying 152 students from the same school arrive at a football stadium. The buses carry (respectively) 39, 33, 46, and 34 students. One of the 152 students is randomly chosen. Let X denote the number of students who were on the bus of the selected student. One of the four bus drivers is also randomly chosen. Let Y be the number of students who were on that driver's bus.

(a) Which do you think is larger, $E[X]$ or $E[Y]$?
(b) Find $E[X]$ and $E[Y]$.

Exercise 1.9 Four buses carrying 152 students from the same school arrive at a football stadium. The buses carry (respectively) 39, 33, 46, and 34 students. One of the 152 students is randomly chosen. Let X denote the number of students who were on the bus of the selected student. One of the four bus drivers is also randomly chosen. Let Y be the number of students who were on that driver's bus.

(a) Which do you think is larger, $E[X]$ or $E[Y]$?
(b) Find $E[X]$ and $E[Y]$.

(b) $E(Y) = \frac{152}{4} = 38$

$$X = \begin{cases} 39 & 39/152 \\ 33 & 33/152 \\ 46 & 46/152 \\ 34 & 34/152 \end{cases}$$

$$E(X) = \frac{39^2 + 33^2 + 46^2 + 34^2}{152}$$

$$= \frac{1521 + 1089 + 2116 + 1156}{152}$$

$$= 38.70$$

$\underline{E(X) > E(Y)}$

(a)

Exercise 1.10 Two players play a tennis match, which ends when one of the players has won two sets. Suppose that each set is equally likely to be won by either player, and that the results from different sets are independent. Find (a) the expected value and (b) the variance of the number of sets played.

(a) X be no. of sets played.

$$X = \begin{cases} 2 & p = 1/2 \\ 3 & 1-p = 1/2 \end{cases}$$

$P(X=2) = AA$ or BB $= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$

$P(X=3) = ABA$ or BAA or BAB or ABB $= 4 \times \frac{1}{8} = \frac{1}{2}$

$E(X) = \frac{2+3}{2} = \underline{\underline{2.5}}$

(b) $Var(X) = E((X - E(X))^2) = p(1-p) = \frac{1}{2} \times \frac{1}{2} = 1/4 = \underline{\underline{0.25}}$

Exercise 1.11 Verify that

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

Hint: Starting with the definition

$$\mathrm{Var}(X) = E[(X - E[X])^2],$$

square the expression on the right side; then use the fact that the expected value of a sum of random variables is equal to the sum of their expectations.

$$Var(X) = E((X - E(X))^2) = E(X^2 - 2E(X)\cdot X + E^2(X))$$
$$= E(X^2) - 2E^2(X) + E(X) = \underline{\underline{E(X^2) - E^2(X)}}$$

Exercise 1.12 A lawyer must decide whether to charge a fixed fee of \$5,000 or take a contingency fee of \$25,000 if she wins the case (and 0 if she loses). She estimates that her probability of winning is .30. Determine the mean and standard deviation of her fee if

(a) she takes the fixed fee;
(b) she takes the contingency fee.

Let X be the fee.

(a) 5000, 0

(b) $X = \begin{cases} 25000 & 0.3 \\ 0 & 0.7 \end{cases}$ $\qquad E(X) = \underline{\underline{7500}}$

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

$X^2 = \begin{cases} 625 \times 10^6 & 0.3 \\ 0 & 0.7 \end{cases}$

$V(X) = \frac{625 \times 10^{5} \times 3}{\cancel{10}} - 625 \times 9 \times 10^4$

$= 30 \times 625 \times 10^4 - 9 \times (625 \times 10^4)$

$V(X) = 21 \times 625 \times 10^4$

$s.d = 2500\sqrt{21} = \underline{\underline{11456.44}}$

Exercise 1.13 Let $X_1, \ldots, X_n$ be independent random variables, all having the same distribution with expected value μ and variance σ^2. The random variable $\bar{X}$, defined as the arithmetic average of these variables, is called the *sample mean*. That is, the sample mean is given by

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}.$$

(a) Show that $E[\bar{X}] = \mu$.
(b) Show that $\text{Var}(\bar{X}) = \sigma^2/n$.

The random variable S^2, defined by

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1},$$

is called the *sample variance*.

(c) Show that $\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$.
(d) Show that $E[S^2] = \sigma^2$.

(a) $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) = \cancel{\frac{1}{n}} \times \cancel{n} \times \mu = \mu$

(b) $Var(\bar{x}) = Var\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) = \frac{\cancel{n}}{n^{\cancel{2}}} \times \sigma^2 = \frac{\sigma^2}{n}$

(c) $S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$ (Sample variance)

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} (X_i^2 - 2\bar{X} X_i + \bar{X}^2)$$
$$= \sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$$

(d) $E(S^2) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} (X_i - \bar{X})^2\right)$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$
$$= \frac{1}{n-1} \left[n E(X_i^2) - n E(\bar{X}^2)\right]$$
$$= n \left[V(X_i) + E^2(X_i) - V(\bar{X}) - E^2(\bar{X})\right]$$

$$= \frac{n}{n-1}\left[V(X_i) + E^2(X_i) - V(\bar{X}) - E^2(\bar{X})\right]$$

$$= \frac{n}{n-1}\left[\sigma^2 + \cancel{\mu^2} - \frac{\sigma^2}{n} - \cancel{\mu^2}\right] = \cancel{\left(\frac{n}{n-1}\right)} \sigma^2 \cancel{\left(\frac{n-1}{n}\right)}$$

$$= \sigma^2 \quad \text{HP//}$$

Exercise 1.14 Verify that

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y].$$

$$\text{Cov}(X,Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

$$= E(XY - XE(Y) - YE(X) + E(X)E(Y))$$

$$= E(XY) - 2E(X)E(Y) + E(X)E(Y)$$

$$= E(XY) - E(X)E(Y)$$

Exercise 1.15 Prove:

(a) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
(b) $\text{Cov}(X, X) = \text{Var}(X)$;
(c) $\text{Cov}(cX, Y) = c\,\text{Cov}(X, Y)$;
(d) $\text{Cov}(c, Y) = 0$.

(a) $\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = E(YX) - E(Y)E(X)$
$= \text{Cov}(Y,X)$

(b) $\text{Cov}(X,X) = E(X^2) - E^2(X) = \text{Var}(X)$

(c) $\text{Cov}(cX,Y) = E(cXY) - E(cX)E(Y) = c(E(XY) - E(X)E(Y))$
$= c\,\text{Cov}(X,Y)$

(d) $\text{Cov}(c,Y) = E(cY) - cE(Y) = 0$//

Exercise 1.16 If U and V are independent random variables, both having variance 1, find $\mathrm{Cov}(X, Y)$ when

$$X = aU + bV, \qquad Y = cU + dV.$$

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(aU + bV, cU + dV) = a\,\mathrm{Cov}(U, cU + dV) + b\,\mathrm{Cov}(V, cU + dV)$$

$$= ac\,\mathrm{Var}(U) + ad\,\mathrm{Cov}(U,V) + bc\,\mathrm{Cov}(U,V) + bd\,\mathrm{Var}(V)$$

$$= \underline{ac + bd}$$

Exercise 1.17 If $\mathrm{Cov}(X_i, X_j) = ij$, find

(a) $\mathrm{Cov}(X_1 + X_2,\ X_3 + X_4)$;
(b) $\mathrm{Cov}(X_1 + X_2 + X_3,\ X_2 + X_3 + X_4)$.

(a) $\mathrm{Cov}(X_1 + X_2, X_3 + X_4) = \mathrm{Cov}(X_1, X_3 + X_4) + \mathrm{Cov}(X_2, X_3 + X_4)$

$= \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_1, X_4) + \mathrm{Cov}(X_2, X_3) + \mathrm{Cov}(X_2, X_4)$

$= 3 + 4 + 6 + 8 = 21$ //

(b) $\mathrm{Cov}(X_1 + X_2 + X_3, X_2 + X_3 + X_4) = \underline{2+3+4} + \underline{4+6+8} + \underline{6+9+12}$

$= 9 + \underline{18 + 27}$

$= 54$ //

★ **Exercise 1.18** Suppose that – in any given time period – a certain stock is equally likely to go up 1 unit or down 1 unit, and that the outcomes of different periods are independent. Let X be the amount the stock goes up (either 1 or -1) in the first period, and let Y be the cumulative amount it goes up in the first three periods. Find the correlation between X and Y.

$$X = \begin{cases} 1 & 1/2 \\ -1 & 1/2 \end{cases}$$

$$Y = X_1 + X_2 + X_3 \Rightarrow Y = \begin{cases} -3 & 1/8 \\ -1 & 3/8 \\ 1 & 3/8 \\ 3 & 1/8 \end{cases}$$

$E(X) = 0 \Rightarrow E(Y) = 0$

$V(X) = E(X^2) = \frac{1}{2}1 + \frac{1}{2}1 = 1 \Rightarrow V(Y) = 3 \cdot 1 = 3$ //

$E(X)$ ...

$V(X) = E(X^2) = \frac{1}{2}1 + \frac{1}{2}1 = 1 \Rightarrow V(Y) = 3\cdot 1 = 3$

$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} \quad \frac{Cov(X_1, X_1+X_2+X_3)}{\sqrt{3}} = \frac{1+0+0}{\sqrt{3}} = \frac{1}{\sqrt{3}}$

Exercise 1.19 Can you construct a pair of random variables such that $\text{Var}(X) = \text{Var}(Y) = 1$ and $\text{Cov}(X, Y) = 2$?

$Cov(X,Y) = E(XY) - E(X)E(Y)$

$X = \begin{cases} 1 & 1/2 \\ -1 & 1/2 \end{cases} \qquad Y = \begin{cases} t & 1/2 \\ -t & 1/2 \end{cases}$

$Y = tX \qquad Cov(X,Y) = t\,Var(X) = t = 2$

$Var(Y) = 4$ ✗

$\therefore X \sim Rad(1/2)$ & $Y \sim 2\,Rad(1/2)$

Possible: $\rho(X,Y) = \frac{2}{\sqrt{1\cdot 1}} = 2 \quad \therefore$ Not Possible

Exercise 1.20 If Y is a random variable and h a function, then $h(Y)$ is also a random variable. If the set of distinct possible values of $h(Y)$ are $\{h_i, i \geq 1\}$, then by the definition of expected value, we have that $E[h(Y)] = \sum_i h_i P(h(Y) = h_i)$. On the other hand, because $h(Y)$ is equal to $h(y)$ when $Y = y$, it is intuitive that

$$E[h(Y)] = \sum_y h(y)P(Y = y)$$

Verify that the preceding equation is valid.

$E(h(Y)) = \sum_i h_i P(h(Y) = h_i)$

$\{h_i\}$ are possible values of $h(Y)$

$S = \{y \mid h(y) = h_i \quad i \geq 1\}$

$$E(h(Y)) = \sum_{y \in S} h(y)\ P(Y = y)$$

Exercise 1.21 The *distribution function* $F(x)$ of the random variable X is defined by

$$F(x) = P(X \leq x)$$

If X takes on one of the values $1, 2, \ldots,$ and F is a known function, how would you obtain $P(X = i)$?

$$P(X = i) = P(X \leq i) - P(X < i)$$
$$= P(X \leq i) - P(X \leq i-1)$$
$$= \underline{F(i) - F(i-1)}$$