

1. (15) Define the following terms:

(a) The distribution function of a random variable.

Let X be a random variable. Then the distribution function $F(x)$ of X equals

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

(b) The covariance of two random variables.

The covariance of the random variables X and Y is

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$

(c) The conditional expectation of X given Y , $E[X|Y]$.

$$E[X|Y] = \sum_x xp_{X|Y}(x|y),$$

where $p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)}$. This formula assumes that X and Y are discrete random variables.

2. (5) An urn contains 15 red balls and 5 blue balls. If two balls are picked from the urn without replacement, what is the probability that they are both blue?

$$\frac{\binom{5}{2}}{\binom{20}{2}} = \frac{1}{19} \approx 0.0526$$

3. (15) Stores A , B , and C have 50, 75, and 100 employees and, respectively, 50, 60, 70 percent of these are women. Resignations are equally likely among all employees, regardless of sex. One employee resigns, and this is a woman. What is the probability that she works in store C ?

The easiest way to work this problem is to say that there are a total of 140 women, 70 of whom work in store C . Thus, the probability that a woman comes from store C is

$$\frac{70}{140} = \frac{1}{2}$$

The harder way is to use Bayes' ideas

$$\begin{aligned} P\{C|w\} &= \frac{P\{C, w\}}{P\{w\}} = \frac{P\{w|C\}P\{C\}}{P\{w|A\}P\{A\} + P\{w|B\}P\{B\} + P\{w|C\}P\{C\}} \\ &= \frac{(7/10)(4/9)}{(1/2)(2/9) + (3/5)(3/9) + (7/10)(4/9)} \\ &= \frac{1}{2} \end{aligned}$$

4. (15) The table below gives the values of the joint probability mass function of two discrete random variables X and Y . Note that the values of X are arranged vertically and those of Y are horizontal. That is, $p(3, -1) = 0$, where $X = 3$ and $Y = -1$.

| $X \setminus Y$ | -1 | 0 | 1 | 2 |
|-----------------|------|------|------|------|
| 1 | 1/10 | 1/10 | 1/20 | 1/6 |
| 2 | 1/50 | 1/10 | 0 | 1/50 |
| 3 | 0 | 1/50 | 1/5 | 1/30 |
| 4 | 1/20 | 1/10 | 1/25 | 0 |

(a) $p(2|Y = 1) = ?$

$$p(2|Y = 1) = \frac{p(2, 1)}{p_Y(1)} = \frac{0}{29/100} = 0$$

(b) $E[X|Y = 1] = ?$

$$\begin{aligned} E[X|Y = 1] &= 1p_{X|Y}(1|1) + 2p_{X|Y}(2|1) + 3p_{X|Y}(3|1) + 4p_{X|Y}(4|1) \\ &= \frac{1/20}{29/100} + 2(0) + 3\frac{1/5}{29/100} + 4\frac{1/25}{29/100} \\ &= \frac{81}{29} \end{aligned}$$

5. (15) In a game a fair die is rolled, and the amount of money won or lost is determined by the number rolled. The amounts are given in the table below

| Number | 1 | 2 | 3 | 4 | 5 | 6 |
|----------|---|---|---|----|---|---|
| Winnings | ? | 3 | 1 | -8 | 1 | 5 |

Let X be the random variable, which denotes the amount of money won or lost.

- (a) If a one is rolled, what should the winnings be so that the expected value of X is -1 . (House advantage)

$$E[X] = \frac{1}{6}(c + 3 + 1 - 8 + 1 + 5) = -1$$

implies that $c = -8$, where c represents the winings if a one is rolled.

- (b) Using the value you found in part a determine the variance of X .

$$\begin{aligned} \text{Var}(X) &= E[X^2] - (E[X])^2 \\ &= \frac{1}{6}(64 + 9 + 1 + 64 + 1 + 25) - (-1)^2 \\ &= \frac{79}{3} \end{aligned}$$

6. (20) A man and a woman arrive at an airport sometime between noon and 1:00 p.m. The woman's arrival time expressed as the number of minutes past noon is uniformly distributed over this 60 minute time period, and is independent of the man's arrival time. If f denotes the probability density function of the man's arrival time in the number of minutes past noon, then

$$f(t) = \begin{cases} \frac{1}{120}, & 0 \leq t < 20 \\ \frac{1}{30}, & 20 \leq t < 40 \\ \frac{1}{120}, & 40 \leq t \leq 60 \\ 0, & \text{otherwise} \end{cases}$$

- (a) What is the probability that the man arrives first?

If we let X denote the random variable which denotes the man's arrival time, and Y denote the woman's arrival time, then the probability that the man arrives first is

$$\begin{aligned} P\{X < Y\} &= \int_0^{60} dx \int_x^{60} \frac{f}{60} dy \\ &= \int_0^{20} dx \int_x^{60} \frac{1}{60 * 120} dy + \int_{20}^{40} dx \int_x^{60} \frac{1}{60 * 30} dy + \int_{40}^{60} dx \int_x^{60} \frac{1}{60 * 120} dy \\ &= \frac{5}{36} + \frac{1}{3} + \frac{1}{36} = \frac{1}{2} \end{aligned}$$

- (b) What is the probability that the man arrives first given that the woman arrives after 12:30?

$$\begin{aligned} P\{X < Y | Y > 30\} &= \frac{P\{X < Y, Y > 30\}}{P\{Y > 30\}} \\ &= \frac{\int_{30}^{40} dx \int_x^{60} \frac{1}{60 * 30} dy + \int_{40}^{60} dx \int_x^{60} \frac{1}{60 * 120} dy}{1/2} \\ &= \frac{5/36 + 1/36}{1/2} = \frac{1}{3} \end{aligned}$$

7. (20) One of the more remarkable things about normal random variables is that sums of independent ones are also normal. As a first step to this we saw that if X is a normally distributed random variable with mean μ and variance σ^2 , then the random variable $Y = aX + b$ is also normal with mean $a\mu + b$ and variance $a^2\sigma^2$.

- (a) What is the density function of a normally distributed random variable with mean 5 and variance 36?

$$\frac{1}{6\sqrt{2\pi}} e^{-\frac{(x-5)^2}{72}}$$

- (b) Verify that $Y = aX + b$ is also normal distributed with mean $a\mu + b$ and variance $a^2\sigma^2$, where μ is the mean of X and σ^2 is the variance of X .

The problem is not to show that the mean and variance are what is expected, but to show that Y is normally distributed. To that end let $G(y)$ denote the distribution function of Y , then we have

$$\begin{aligned} G(y) &= P\{Y \leq y\} = P\{aX + b \leq y\} \\ &= P\left\{X \leq \frac{y-b}{a}\right\} \text{ assume that } a > 0 \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{y-b}{a}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \end{aligned}$$

To get the density function, $g(y)$ of Y , we differentiate the distribution function of Y

$$\begin{aligned} g(y) &= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{((y-b)/a - \mu)^2}{2\sigma^2}} \\ &= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{(y - (a\mu + b))^2}{2a^2\sigma^2}} \end{aligned}$$

Thus, we see that Y is normally distributed with mean $a\mu + b$ and variance $a^2\sigma^2$.

8. (15) Suppose that a coin is flipped 3200 times and the probability that it comes up heads is $1/4$. Find, by using a normally distributed random variable, an approximation to the probability that there are between 740 and 750 heads.

If X is the random variable which counts the number of heads flipped then the expected value of X and its variance equal

$$\begin{aligned} E[X] &= 3200 \left(\frac{1}{4}\right) = 800 \\ \text{var}(X) &= 3200 \left(\frac{1}{4}\right) \left(\frac{3}{4}\right) = 600 \end{aligned}$$

Let N denote a random variable which is normally distributed with parameters 800 and 600. Then

$$\begin{aligned}
 P\{740 \leq X \leq 750\} &\approx P\{739.5 \leq N \leq 750.5\} \\
 &= P\left\{\frac{739.5 - 800}{10\sqrt{6}} \leq \frac{N - 800}{10\sqrt{6}} \leq \frac{750.5 - 800}{10\sqrt{6}}\right\} \\
 &\approx P\{-2.469 \leq Z \leq -2.0208\} \\
 &= P\{2.02 \leq Z \leq 2.47\} \\
 &= 0.9932 - 0.9783 \\
 &= 0.0149
 \end{aligned}$$

9. (15) A model proposed for college football supposes that when two teams with roughly the same record play each other, then the number of points scored in a half by the home team minus the number scored by the visiting team is approximately a normal random variable with mean 5 and variance 8. In addition the model supposes that the point differentials for the two halves are independent. Assuming this model

- (a) What is the probability that the home team wins?

Let H_1 denote the score differential in the first half and H_2 the score differential in the second half. Then $H_1 + H_2$ is a normally distributed random variable with mean 10 and variance 16. The probability that the home team wins is then approximated by

$$\begin{aligned}
 P\left\{H_1 + H_2 > \frac{1}{2}\right\} &= P\left\{\frac{H_1 + H_2 - 10}{4} > -\frac{9.5}{4}\right\} \\
 &= P\{Z > -2.375\} = P\{Z < 2.375\} \\
 &\approx 0.9911
 \end{aligned}$$

- (b) Given that the home team is behind by 3 points at the half, what is the conditional probability that the home team wins?

$$\begin{aligned}
 P\left\{H_1 + H_2 > \frac{1}{2} \mid -3.5 < H_1 < -2.5\right\} &= \frac{P\left\{H_1 + H_2 > \frac{1}{2}, -3.5 < H_1 < -2.5\right\}}{P\{-3.5 < H_1 < -2.5\}} \\
 &= \frac{0.0006}{0.0015} \\
 &= 0.40117
 \end{aligned}$$

The computation of these probabilities was done using Maple.

10. (15) A prisoner is trapped in a cell containing 3 doors. The first door leads to a tunnel that returns him to his cell after 2 days travel. The second leads to a tunnel that returns him to his cell after 4 days travel. The third door leads to freedom after 1 day of travel. If it is assumed that the prisoner will always select doors 1, 2, and 3 with respective probabilities 0.5, 0.3, and 0.2, what is the expected number of days until the prisoner reaches freedom?

$$\begin{aligned} E[X] &= E[E[X|Y]] \\ &= E[X|Y=1]P\{Y=1\} + E[X|Y=2]P\{Y=2\} \\ &\quad + E[X|Y=3]P\{Y=3\} \\ &= (2 + E[X])\frac{1}{2} + (4 + E[X])\frac{3}{10} + 1\left(\frac{2}{10}\right) \end{aligned}$$

Solving the equation above for $E[X]$ we get

$$E[X] = 12$$