STAT451 HW6

Russell Miller

November 17, 2011

4.39 The total number of hours, in units of 100 hours, that a family runs a vacuum cleaner over a period of one year is a random variable X having the density function given in Exercise 4.15 on page 114. Find the variance of X.

$$f(x) = \begin{cases} x, & 0 < x < 1, \\ 2 - x, & 1 \le x < 2, \\ 0, & \text{elsewhere} \end{cases}$$

Exercise 4.15 asked for the average number of hours per year that families run their vacuum cleaners. This is also known as $E[X]$. We found that

$$E[X] = 1$$

The other piece we will need is $E[X^2]$. By the law of the unconscious statistician, we can let $g(X) = X^2$ and find $E[g(X)]$.

$$\begin{aligned}
E[g(X)] = E[X^2] &= \int_0^1 (x^2)x dx + \int_1^2 (x^2)(2-x) dx \\
&= \int_0^1 x^3 dx + \int_1^2 (2x^2 - x^3) dx \\
&= \left.\frac{x^4}{4}\right|_0^1 + \left.\left(\frac{2x^3}{3} - \frac{x^4}{4}\right)\right|_1^2 \\
&= \frac{1}{4} + \left(\frac{16}{3} - 4\right) - \left(\frac{2}{3} - \frac{1}{4}\right) \\
&= \frac{1}{4} + \frac{4}{3} - \frac{5}{12} \\
&= \frac{7}{6}
\end{aligned}$$

And we have the variance

$$\begin{aligned}
V[X] &= E[X^2] - E[X]^2 \\
&= \frac{7}{6} - 1 \\
&= \boxed{\frac{1}{6}}
\end{aligned}$$

4.41 Find the standard deviation of the random variable $g(X) = (2X+1)^2$ in Exercise 4.17 on page 114.

x	-3	6	9
$f(x)$	1/6	1/2	1/3

Exercise 4.17 asked for $\mu_{g(x)}$. We found that it was

$$\begin{aligned}\mu_{g(x)} = E[g(X)] &= \sum_{x=-3}^{9} (2x+1)^2 f(x) \\ &= (25)\left(\frac{1}{6}\right) + (169)\left(\frac{1}{2}\right) + (361)\left(\frac{1}{3}\right) \\ &= \frac{25+507+722}{6} \\ &= 209\end{aligned}$$

We're also going to need $E[(g(X))^2]$.

$$\begin{aligned}E[(g(X))^2] &= (625)\left(\frac{1}{6}\right) + (28561)\left(\frac{1}{2}\right) + (130321)\left(\frac{1}{3}\right) \\ &= 57825\end{aligned}$$

Now we have the variance

$$\begin{aligned}\sigma^2 = V[g(X)] &= E[(g(X))^2] - E[g(X)]^2 \\ &= 57825 - 209^2 \\ &= 14144\end{aligned}$$

And the standard deviation is

$$\begin{aligned}\sigma &= \sqrt{\sigma^2} \\ &= \boxed{118.9}\end{aligned}$$

4.42 Using the results of Exercise 4.21 on page 114, find the variance of $g(X) = X^2$, where X is a random variable having the density function given in Exercise 4.12 on page 113.

$$f(x) = \begin{cases} 2(1-x), & 0 < x < 1, \\ 0, & \text{elsewhere} \end{cases}$$

The solution of 4.21 was the average, or $E[g(X)]$.

$$\begin{aligned}E[g(X)] = E(X^2) &= \int_0^1 (x^2)2(1-x)dx \\ &= \int_0^1 2(x^2 - x^3)dx \\ &= \frac{1}{6}\end{aligned}$$

We need to square that.

$$\begin{aligned}E[g(X)^2] = E(X^4) &= \int_0^1 (x^4)2(1-x)dx \\ &= \int_0^1 2(x^4 - x^5)dx \\ &= \frac{1}{15}\end{aligned}$$

Now we can calculate variance.

$$\begin{aligned} V[g(X)] &= E[g(X)^2] - E[g(X)]^2 \\ &= \frac{1}{15} - \left(\frac{1}{6}\right)^2 \\ &= \boxed{\frac{7}{180} \approx .0389} \end{aligned}$$

4.44 Find the covariance of the random variables X and Y of Exercise 3.39 on page 101.
Here's the table from the joint probability function $f(x,y)$, adding in the marginal values.

$f(x,y)$	x = 0	1	2	3	$f_2(y)$
y = 0	0	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	$\frac{15}{70}$
1	$\frac{2}{70}$	$\frac{18}{70}$	$\frac{18}{70}$	$\frac{2}{70}$	$\frac{40}{70}$
2	$\frac{3}{70}$	$\frac{9}{70}$	$\frac{3}{70}$	0	$\frac{15}{70}$
$f_1(x)$	$\frac{5}{70}$	$\frac{30}{70}$	$\frac{30}{70}$	$\frac{5}{70}$	1

Now the first thing we'll need is μ_X

$$\mu_X = \sum_{x=0}^{3} x f_1(x) = (0)\left(\frac{5}{70}\right) + (1)\left(\frac{30}{70}\right) + (2)\left(\frac{30}{70}\right) + (3)\left(\frac{5}{70}\right) = \frac{3}{2}$$

And, of course, μ_Y

$$\mu_Y = \sum_{y=0}^{2} y f_2(y) = (0)\left(\frac{15}{70}\right) + (1)\left(\frac{40}{70}\right) + (2)\left(\frac{15}{70}\right) = 1$$

We also need $E[XY]$

$$\begin{aligned} E[XY] = \sum_{x=0}^{3}\sum_{y=0}^{2} xyf(x,y) &= (0)(1)f(0,1) + (0)(2)f(0,2) + (1)(0)f(1,0) + \\ & (1)(1)f(1,1) + (1)(2)f(1,2) + (2)(0)f(2,0) + \\ & (2)(1)f(2,1) + (2)(2)f(2,2) + (3)(0)f(3,0) + \\ & (3)(1)f(3,1) = \frac{9}{7} \end{aligned}$$

And we put it all together to get the covariance.

$$\sigma_{XY} = E[XY] - \mu_X\mu_Y = \frac{9}{7} - \left(\frac{3}{2}\right)(1) = \boxed{-\frac{3}{14}}$$

4.45 Find the covariance of the random variables X and Y of Exercise 3.49 on page 102.
Alright, here goes $f(x,y)$, $f_1(x)$, $f_2(y)$, μ_X, μ_Y, $E[XY]$.

$f(x,y)$	x = 1	2	3	$f_2(y)$
y = 1	0.05	0.05	0.1	0.2
2	0.05	0.1	0.35	0.5
3	0	0.2	0.1	0.3
$f_1(x)$	0.1	0.35	0.55	1

$$\begin{aligned}
\mu_X &= \sum_{x=1}^{3} x f_1(x) = (1)(.1) + (2)(.35) + (3)(.55) = 2.45 \\
\mu_Y &= \sum_{y=1}^{3} y f_2(y) = (1)(.2) + (2)(.5) + (3)(.5) = 2.1 \\
E[XY] &= \sum_{x=1}^{3}\sum_{y=1}^{3} xyf(x,y) = (1)(1)f(1,1) + (1)(2)f(1,2) + (1)(3)f(1,3) + \\
&\quad (2)(1)f(2,1) + (2)(2)f(2,2) + (2)(3)f(2,3) + \\
&\quad (3)(1)f(3,1) + (3)(2)f(3,2) + (3)(3)f(3,3) \\
&= 5.15 \\
\sigma_{XY} &= E[XY] - \mu_X\mu_Y = 5.15 - (2.45)(2.1) = \boxed{0.005}
\end{aligned}$$

4.48 Given a random variable X, with standard deviation σ_X and a random variable $Y = a + bX$, show that if $b < 0$, the correlation coefficient $\rho_{XY} = -1$, and if $b > 0$, $\rho_{XY} = 1$.

$$\begin{aligned}
\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} &= \frac{E[XY] - E[X]E[Y]}{\sqrt{E[X^2] - E[X]}\sqrt{E[Y^2] - E[Y]}} \\
&= \frac{E[X(a+bX)] - E[X]E[a+bX]}{\sqrt{E[X^2] - E[X]}\sqrt{E[(a+bX)^2] - E[a+bX]}} \\
&= \frac{E[Xa + bX^2] - bE[X]^2}{V[X]^2\sqrt{E[a^2 + 2abX + b^2X^2] - bE[X]}} \\
&= \frac{aE[X] + bE[X^2] - bE[X]^2}{V[X]^2\sqrt{2abE[X] + b^2E[X^2] - bE[X]}} \\
&= \frac{aE[X] + bV[X]}{V[X]^2\sqrt{b(2aE[X] + bE[X^2] - E[X])}}
\end{aligned}$$

I'm dizzy, I give up.

4.51 Referring to Exercise 4.35 on page 122, find the mean and variance of the discrete random variable $Z = 3X - 2$, when X represents the number of errors per 100 lines of code.

x	2	3	4	5	6
$f(x)$	0.01	0.25	0.4	0.3	0.04

$$\begin{aligned}
\mu_Z = E[3X-2] &= \sum_{x=2}^{6}(3x-2)f(x) \\
&= (4\times 0.01) + (7\times 0.25) + (10\times 0.4) + (13\times 0.3) + (16\times 0.04) \\
&= \boxed{10.33} \leftarrow \text{ mean} \\
E[(3X-2)^2] &= \sum_{x=2}^{6}(3x-2)^2 f(x) \\
&= (16\times 0.01) + (49\times 0.25) + (100\times 0.4) + (169\times 0.3) + (256\times 0.04) \\
&= 113.35 \\
\sigma_Z^2 = E[Z^2] - E[Z]^2 &= 113.35 - 10.33^2 \\
&= \boxed{6.6411} \leftarrow \text{ variance}
\end{aligned}$$

4.59 Use Theorem 4.7 to evaluate $E[2XY^2 - X^2Y]$ for the joint probability distribution shown in Table 3.1 on page 92.

Table 3.1 Joint Probability Distribution

		x			
$f(x,y)$		0	1	2	$f_2(y)$
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
$f_1(x)$		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

By Theorem 4.7 we see that we can rewrite this as

$$E[2XY^2 - X^2Y] = E[2XY^2] - E[X^2Y]$$

First we'll find $E[2XY^2]$

$$E[2XY^2] = \sum_{x=0}^{2}\sum_{y=0}^{2} 2xy^2 f(x,y) = 2(0)(0^2)\left(\frac{3}{28}\right) + 2(0)(1^2)\left(\frac{3}{14}\right) +$$

$$2(0)(2^2)\left(\frac{1}{28}\right) + 2(1)(0^2)\left(\frac{9}{28}\right) + 2(1)(1^2)\left(\frac{3}{14}\right) + 2(2)(0^2)\left(\frac{3}{28}\right) = \frac{3}{7}$$

Next $E[X^2Y]$

$$E[X^2Y] = \sum_{x=0}^{2}\sum_{y=0}^{2} x^2 y f(x,y) = (0^2)(0)\left(\frac{3}{28}\right) + (0^2)(1)\left(\frac{3}{14}\right) +$$

$$(0^2)(2)\left(\frac{1}{28}\right) + (1^2)(0)\left(\frac{9}{28}\right) + (1^2)(1)\left(\frac{3}{14}\right) + (2^2)(0)\left(\frac{3}{28}\right) = \frac{3}{14}$$

And the grande finale

$$E[2XY^2] - E[X^2Y] = \frac{3}{7} - \frac{3}{14} = \boxed{\frac{3}{14}}$$

4.60 Seventy new jobs are opening up at an automobile manufacturing plant, but 1000 applicants show up for the 70 positions. To select the best 70 from among the applicants, the company gives a test that covers mechanical skill, manual dexterity, and mathematical ability. The mean grade on this test turns out to be 60, and the scores have a standard deviation 6. Can a person who has an 84 score count on getting one of the jobs? [*Hint*: Use Chebyshev's theorem.] Assume that the distribution is symmetric about the mean.

We can see that someone who scores an 84 is 4 standard deviations above the mean grade. By Chebyshev's Theorem we know that

$$\begin{aligned} P(\mu - k\sigma < X < \mu + k\sigma) &\geq 1 - \frac{1}{k^2} \\ P(60 - 4\times 6 < X < 60 + 4\times 6) &\geq 1 - \frac{1}{4^2} \end{aligned}$$

But the part we care about is

$$P(X < 84) \geq 0.9375$$

So scoring an 84 means only 0.0625 of the scores are as good. So because the top 0.07 of the applicants are going to get a job, you could certainly count on getting one.

4.63 Suppose that you roll a fair 10-sided die (0,1,2,...,9) 500 times. Using Chebyshev's theorem, compute the probability that the sample mean, X is between 4 and 5.

I've never seen a die with a zero, but any way...

$$\begin{aligned}
\mu &= \sum_{x=0}^{9} x f(x) = (45)\left(\frac{1}{10}\right) = 4.5 \\
E[X^2] &= \sum_{x=0}^{9} x^2 f(x) = (285)\left(\frac{1}{10}\right) = 28.5 \\
\sigma^2 &= 28.5 - (4.5)^2 = 8.25 \\
\sigma &= \sqrt{\sigma} \approx 2.87
\end{aligned}$$

Now we're ready to use Chebyshev's theorem.

$$\begin{aligned}
P(\mu - k\sigma < X < \mu + k\sigma) &\geq 1 - \frac{1}{k^2} \\
P(4.5 - k2.87 < X < 4.5 + k2.87) &\geq 1 - \frac{1}{k^2}
\end{aligned}$$

In order to set up the inequality, we let $k = 0.1742$.

$$\begin{aligned}
P(4 < X < 5) &\geq 1 - \frac{1}{0.0303} \\
&\geq -32.003
\end{aligned}$$

Horrible bound. I did something wrong.

4.67 A random variable X has a mean $\mu = 10$ and a variance $\sigma^2 = 4$. Using Chebyshev's theorem, find

a. $P(|X - 10| \geq 3)$

$$P(|X-10| \geq 3) = 1 - P(|X-10| < 3) = 1 - P(-3 < X - 10 < 3) = 1 - P(10 - 3 < X < 10 + 3)$$

$k\sigma = 3$, so $k = 3/2$

$$1 - P(10 - (3/2)(2) < X < 10 + (3/2)(2)) \boxed{\leq 4/9}$$

b. $P(|X - 10| < 3)$

$$\begin{aligned}
P(10 - (3/2)(2) < X < 10 + (3/2)(2)) &\geq 1 - \frac{4}{9} \\
&\boxed{\geq \frac{5}{9}}
\end{aligned}$$

c. $P(5 < X < 15)$

Given $\mu = 10$ and $\sigma = 2$, $10 - 2k = 5$ and $10 + 2k = 15$. Thus $k = 5/2$.

$$\begin{aligned}
P(10 - (5/2)(2) < X < 10 + (5/2)(2)) &\geq 1 - (4/25) \\
&\boxed{\geq \frac{21}{25}}
\end{aligned}$$

d. the value of the constant c such that $P(|X - 10| \geq c) \leq 0.04$

$$\begin{aligned}
P(|X-10| \geq c) &= 1 - P(|X-10| < c) = 1 - P(-c < X - 10 < c) = 1 - P(10 - c < X < 10 + c) \\
0.04 &= \frac{1}{k^2} \\
k &= 5 \\
c &= 2k
\end{aligned}$$

$\boxed{c = 10}$

4.69 Let X represent the number that occurs when a red die is tossed and Y the number that occurs when a green die is tossed. Find

a. $E[X+Y]$

$$\begin{aligned} E[X+Y] = E[X] + E[Y] &= \sum_{x=1}^{6} xf(x) + \sum_{y=1}^{6} yf(y) \\ &= (21)\left(\frac{1}{6}\right) + (21)\left(\frac{1}{6}\right) \\ &= \boxed{7} \end{aligned}$$

b. $E[X-Y]$

$$\begin{aligned} E[X-Y] = E[X] - E[Y] &= \sum_{x=1}^{6} xf(x) - \sum_{y=1}^{6} yf(y) \\ &= (21)\left(\frac{1}{6}\right) - (21)\left(\frac{1}{6}\right) \\ &= \boxed{0} \end{aligned}$$

c. $E[XY]$

$$\begin{aligned} E[XY] &= \sum_{x=1}^{6}\sum_{y=1}^{6} xyf(x,y) \\ &= (21)(21)\left(\frac{1}{36}\right) \\ &= \boxed{12.25} \end{aligned}$$

4.73 Consider a random variable X with density function

$$f(x) = \begin{cases} \frac{1}{5}, & 0 \le x \le 5, \\ 0, & \text{elsewhere.} \end{cases}$$

a. Find $\mu = E[X]$ and $\sigma^2 = E[(X-\mu)^2]$.

$$\begin{aligned} E[X] &= \int_0^5 xf(x)dx = \int_0^5 \frac{x}{5}dx = \boxed{\frac{5}{2}} \\ E[(X-\mu)^2] &= \int_0^5 (x-\mu)^2 f(x)dx = \int_0^5 \left(x - \frac{5}{2}\right)^2 \left(\frac{1}{5}\right) dx = \boxed{\frac{25}{12}} \end{aligned}$$

b. Demonstrate that Chebyshev's theorem holds for $k=2$ and $k=3$.
When $k=2$

$$\begin{aligned} P\left(\frac{5}{2} - 2(2.89) < X < \frac{5}{2} + 2(2.89)\right) &\ge \frac{3}{4} \\ P(-.387 < X < 5.387) &\ge \frac{3}{4} \end{aligned}$$

This holds because $f(x)$ is nonzero when $0 < X < 5$, which lies within this range.
When $k=3$

$$\begin{aligned} P\left(\frac{5}{2} - 3(2.89) < X < \frac{5}{2} + 3(2.89)\right) &\ge \frac{8}{9} \\ P(-1.83 < X < 6.83) &\ge \frac{8}{9} \end{aligned}$$

This holds because $f(x)$ is nonzero when $0 < X < 5$, which lies within this range.

4.77 The length of time Y in minutes required to generate a human reflex to tear gas has density function

$$f(y) = \begin{cases} \frac{1}{4}e^{-y/4}, & 0 \le y < \infty \\ 0, & \text{elsewhere.} \end{cases}$$

a. What is the mean time to reflex?

$$E[Y] = \int_0^\infty y\frac{1}{4}e^{-y/4}dy = \boxed{4}$$

b. Find $E[Y^2]$ and $V[Y]$.

$$E[Y^2] = \int_0^\infty y^2\frac{1}{4}e^{-y/4}dy = \boxed{32}$$
$$V[Y] = E[Y^2] - E[Y]^2 = 32 - 16 = \boxed{16}$$

4.81 Prove Chebyshev's theorem when X is a discrete random variable.
We can write variance as

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) \tag{1}$$
$$= \sum_{x<\mu-k\sigma} (x-\mu)^2 f(x) + \sum_{x>=\mu+k\sigma} (x-\mu)^2 f(x) \tag{2}$$
$$\ge \sum_{x<\mu-k\sigma} k^2\sigma^2 f(x) + \sum_{x>=\mu+k\sigma} k^2\sigma^2 f(x) \tag{3}$$
$$\sum_{x<\mu-k\sigma} f(x) + \sum_{x>=\mu+k\sigma} f(x) \le \frac{1}{k^2} \tag{4}$$
$$P(\mu - k\sigma < X < \mu + k\sigma) = \sum_{x=\mu-k\sigma}^{\mu+k\sigma} f(x) \ge 1 - \frac{1}{k^2} \tag{5}$$

4.91 Consider the joint density function

$$f(x,y) = \begin{cases} \frac{16y}{x^3}, & x > 2, 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Compute the correlation coefficient ρ_{XY}.

First we need the marginal density functions.

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 \frac{16y}{x^3}dy = \left.\frac{8y^2}{x^3}\right|_0^1 \\ &= \frac{8}{x^3}, x > 2 \\ f_2(y) &= \int_{-\infty}^{\infty} f(x,y)dx = \int_2^\infty \frac{16y}{x^3}dx = \left.\frac{-8y}{x^2}\right|_2^\infty \\ &= 2y, 0 < y < 1 \end{aligned}$$

Next we need μ_X and μ_Y.

$$\begin{aligned} \mu_X &= \int_2^\infty \frac{8}{x^2}dx = 4 \\ \mu_Y &= \int_0^1 2y^2 dy = \frac{2}{3} \end{aligned}$$

Now we need $E[XY]$.

$$E[XY] = \int_2^\infty \int_0^1 \frac{16y^2}{x^2} dxdy = \frac{8}{3}$$

And finally we can compute σ_{XY}.

$$\begin{aligned} \sigma_{XY} &= E[XY] - \mu_X \mu_Y = \frac{8}{3} - (4)\left(\frac{2}{3}\right) \\ &= \boxed{0} \end{aligned}$$