CHAPTER 5

Section 5.1

1.

a. $P(X = 1, Y = 1) = p(1,1) = .20$

b. $P(X \le 1 \text{ and } Y \le 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .42$

c. At least one hose is in use at both islands. $P(X \ne 0 \text{ and } Y \ne 0) = p(1,1) + p(1,2) + p(2,1) + p(2,2) = .70$

d. By summing row probabilities, $p_x(x) = .16, .34, .50$ for $x = 0, 1, 2$, and by summing column probabilities, $p_y(y) = .24, .38, .38$ for $y = 0, 1, 2$. $P(X \le 1) = p_x(0) + p_x(1) = .50$

e. $P(0,0) = .10$, but $p_x(0) \cdot p_y(0) = (.16)(.24) = .0384 \ne .10$, so X and Y are not independent.

2.

a.

	p(x,y)	y: 0	1	2	3	4	
	0	.30	.05	.025	.025	.10	.5
x	1	.18	.03	.015	.015	.06	.3
	2	.12	.02	.01	.01	.04	.2
		.6	.1	.05	.05	.2	

b. $P(X \le 1 \text{ and } Y \le 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .56$
$= (.8)(.7) = P(X \le 1) \cdot P(Y \le 1)$

c. $P(X + Y = 0) = P(X = 0 \text{ and } Y = 0) = p(0,0) = .30$

d. $P(X + Y \le 1) = p(0,0) + p(0,1) + p(1,0) = .53$

3.

a. $p(1,1) = .15$, the entry in the 1st row and 1st column of the joint probability table.

b. $P(X_1 = X_2) = p(0,0) + p(1,1) + p(2,2) + p(3,3) = .08+.15+.10+.07 = .40$

c. $A = \{ (x_1, x_2): x_1 \ge 2 + x_2 \} \cup \{ (x_1, x_2): x_2 \ge 2 + x_1 \}$
$P(A) = p(2,0) + p(3,0) + p(4,0) + p(3,1) + p(4,1) + p(4,2) + p(0,2) + p(0,3) + p(1,3) =.22$

d. $P(\text{ exactly } 4) = p(1,3) + p(2,2) + p(3,1) + p(4,0) = .17$
$P(\text{at least } 4) = P(\text{exactly } 4) + p(4,1) + p(4,2) + p(4,3) + p(3,2) + p(3,3) + p(2,3)=.46$

4.

a. $P_1(0) = P(X_1 = 0) = p(0,0) + p(0,1) + p(0,2) + p(0,3) = .19$
$P_1(1) = P(X_1 = 1) = p(1,0) + p(1,1) + p(1,2) + p(1,3) = .30$, etc.

x_1	0	1	2	3	4
$p_1(x_1)$	.19	.30	.25	.14	.12

b. $P_2(0) = P(X_2 = 0) = p(0,0) + p(1,0) + p(2,0) + p(3,0) + p(4,0) = .19$, etc

x_2	0	1	2	3
$p_2(x_2)$	.19	.30	.28	.23

c. $p(4,0) = 0$, yet $p_1(4) = .12 > 0$ and $p_2(0) = .19 > 0$, so $p(x_1 , x_2) \neq p_1(x_1) \cdot p_2(x_2)$ for every (x_1 , x_2), and the two variables are not independent.

5.

a. P(X = 3, Y = 3) = P(3 customers, each with 1 package)
= P(each has 1 package | 3 customers) · P(3 customers)
$= (.6)^3 \cdot (.25) = .054$

b. P(X = 4, Y = 11) = P(total of 11 packages | 4 customers) · P(4 customers)
Given that there are 4 customers, there are 4 different ways to have a total of 11 packages: 3, 3, 3, 2 or 3, 3, 2, 3 or 3, 2, 3 ,3 or 2, 3, 3, 3. Each way has probability $(.1)^3(.3)$, so $p(4, 11) = 4(.1)^3(.3)(.15) = .00018$

6.

a. $p(4,2) = P(Y = 2 \mid X = 4) \cdot P(X = 4) = \left[\binom{4}{2}(.6)^2(.4)^2\right] \cdot (.15) = .0518$

b. $P(X = Y) = p(0,0) + p(1,1) + p(2,2) + p(3,3) + p(4,4) = .1+(.2)(.6) + (.3)(.6)^2 + (.25)(.6)^3 + (.15)(.6)^4 = .4014$

c. p(x,y) = 0 unless y = 0, 1, …, x; x = 0, 1, 2, 3, 4. For any such pair,

p(x,y) = P(Y = y | X = x) · P(X = x) = $\binom{x}{y}(.6)^y(.4)^{x-y} \cdot p_x(x)$

$p_y(4) = p(y = 4) = p(x = 4, y = 4) = p(4,4) = (.6)^4 \cdot (.15) = .0194$

$p_y(3) = p(3,3) + p(4,3) = (.6)^3(.25) + \binom{4}{3}(.6)^3(.4)(.15) = .1058$

$p_y(2) = p(2,2) + p(3,2) + p(4,2) = (.6)^2(.3) + \binom{3}{2}(.6)^2(.4)(.25)$

$+\binom{4}{2}(.6)^2(.4)^2(.15) = .2678$

$p_y(1) = p(1,1) + p(2,1) + p(3,1) + p(4,1) = (.6)(.2) + \binom{2}{1}(.6)(.4)(.3)$

$\binom{3}{1}(.6)(.4)^2(.25) + \binom{4}{1}(.6)(.4)^3(.15) = .3590$

$p_y(0) = 1 – [.3590+.2678+.1058+.0194] = .2480$

7.

a. p(1,1) = .030

b. P(X ≤ 1 and Y ≤ 1 = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .120

c. P(X = 1) = p(1,0) + p(1,1) + p(1,2) = .100; P(Y = 1) = p(0,1) + … + p(5,1) = .300

d. P(overflow) = P(X + 3Y > 5) = 1 – P(X + 3Y ≤ 5) = 1 – P[(X,Y)=(0,0) or …or (5,0) or (0,1) or (1,1) or (2,1)] = 1 - .620 = .380

e. The marginal probabilities for X (row sums from the joint probability table) are $p_x(0)$ = .05, $p_x(1)$ = .10 , $p_x(2)$ = .25, $p_x(3)$ = .30, $p_x(4)$ = .20, $p_x(5)$ = .10; those for Y (column sums) are $p_y(0)$ = .5, $p_y(1)$ = .3, $p_y(2)$ = .2. It is now easily verified that for every (x,y), $p(x,y) = p_x(x) \cdot p_y(y)$, so X and Y are independent.

8.

a. numerator = $\binom{8}{3}\binom{10}{2}\binom{12}{1} = (56)(45)(12) = 30{,}240$

denominator = $\binom{30}{6} = 593{,}775$; p(3,2) = $\dfrac{30{,}240}{593{,}775} = .0509$

b. p(x,y) = $\begin{cases} \dfrac{\binom{8}{x}\binom{10}{y}\binom{12}{6-(x+y)}}{\binom{30}{6}} & x, y\_are\_non-negative \\ & \text{int }egers\_such\_that \\ & 0 \le x + y \le 6 \\ 0 & otherwise \end{cases}$

9.

a. $1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_{20}^{30}\int_{20}^{30} K(x^2 + y^2)dxdy$

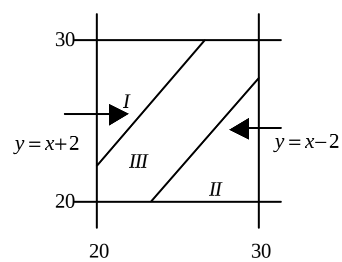
$= K\int_{20}^{30}\int_{20}^{30} x^2 dydx + K\int_{20}^{30}\int_{20}^{30} y^2 dxdy = 10K\int_{20}^{30} x^2 dx + 10K\int_{20}^{30} y^2 dy$

$= 20K \cdot \left(\dfrac{19{,}000}{3}\right) \Rightarrow K = \dfrac{3}{380{,}000}$

b. P(X < 26 and Y < 26) = $\int_{20}^{26}\int_{20}^{26} K(x^2 + y^2)dxdy = 12K\int_{20}^{26} x^2 dx$

$4Kx^3\Big|_{20}^{26} = 38{,}304K = .3024$

c.



P(| X – Y | ≤ 2) = $\iint_{\substack{region \\ III}} f(x,y)dxdy$

$1 - \iint_{I} f(x,y)dxdy - \iint_{II} f(x,y)dxdy$

$1 - \int_{20}^{28}\int_{x+2}^{30} f(x,y)dydx - \int_{22}^{30}\int_{20}^{x-2} f(x,y)dydx$

= (after much algebra) .3593

d. $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{20}^{30} K(x^2+y^2)dy = 10Kx^2 + K\frac{y^3}{3}\Big|_{20}^{30}$

$= 10Kx^2 + .05, \qquad 20 \le x \le 30$

e. $f_y(y)$ is obtained by substituting y for x in (d); clearly $f(x,y) \ne f_x(x) \cdot f_y(y)$, so X and Y are not independent.

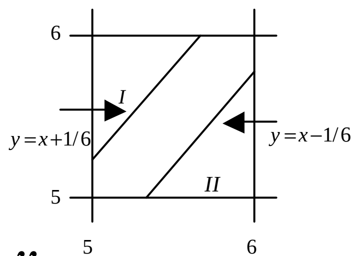
10.

a. $f(x,y) = \begin{cases} 1 & 5 \le x \le 6, 5 \le y \le 6 \\ 0 & otherwise \end{cases}$

since $f_x(x) = 1$, $f_y(y) = 1$ for $5 \le x \le 6$, $5 \le y \le 6$

b. $P(5.25 \le X \le 5.75, 5.25 \le Y \le 5.75) = P(5.25 \le X \le 5.75) \cdot P(5.25 \le Y \le 5.75)$ = (by independence) (.5)(.5) = .25

c.



$P((X,Y) \in A) = \iint_A 1 dxdy$

= area of A = 1 – (area of I + area of II)

$= 1 - \frac{25}{36} = \frac{11}{36} = .306$

11.

a. $p(x,y) = \frac{e^{-l} l^x}{x!} \cdot \frac{e^{-m} m^y}{y!}$ for x = 0, 1, 2, …; y = 0, 1, 2, …

b. $p(0,0) + p(0,1) + p(1,0) = e^{-l-m}[1 + l + m]$

c. $P(X+Y=m) = \sum_{k=0}^{m} P(X=k, Y=m-k) = \sum_{k=0}^{m} e^{-l-m} \frac{l^k}{k!} \frac{m^{m=k}}{(m-k)!}$

$\frac{e^{-(l+m)}}{m!} \sum_{k=0}^{m} \binom{m}{k} l^k m^{m-k} = \frac{e^{-(l+m)}(l+m)^m}{m!}$, so the total # of errors X+Y also has a Poisson distribution with parameter $l + m$.

12.

a. P(X> 3) = $\int_3^\infty \int_0^\infty xe^{-x(1+y)}dydx = \int_3^\infty e^{-x}dx = .050$

b. The marginal pdf of X is $\int_0^\infty xe^{-x(1+y)}dy = e^{-x}$ for $0 \le x$; that of Y is $\int_3^\infty xe^{-x(1+y)}dx = \dfrac{1}{(1+y)^2}$ for $0 \le y$. It is now clear that f(x,y) is not the product of the marginal pdf's, so the two r.v's are not independent.

c. P(at least one exceeds 3) = 1 – P(X ≤ 3 and Y ≤ 3)

$$=1-\int_0^3\int_0^3 xe^{-x(1+y)}dydx = 1-\int_0^3\int_0^3 xe^{-x}e^{-xy}dy$$

$$=1-\int_0^3 e^{-x}(1-e^{-3x})dx = e^{-3}+.25-.25e^{-12} = .300$$

13.

a. $f(x,y) = f_x(x)\cdot f_y(y) = \begin{cases} e^{-x-y} & x \ge 0, y \ge 0 \\ 0 & otherwise \end{cases}$

b. $P(X \le 1 \text{ and } Y \le 1) = P(X \le 1)\cdot P(Y \le 1) = (1 - e^{-1})(1 - e^{-1}) = .400$

c. $P(X + Y \le 2) = \int_0^2\int_0^{2-x} e^{-x-y}dydx = \int_0^2 e^{-x}\left[1-e^{-(2-x)}\right]dx$

$$= \int_0^2 (e^{-x}-e^{-2})dx = 1-e^{-2}-2e^{-2} = .594$$

d. $P(X + Y \le 1) = \int_0^1 e^{-x}\left[1-e^{-(1-x)}\right]dx = 1-2e^{-1} = .264$,

so $P(1 \le X + Y \le 2) = P(X + Y \le 2) - P(X + Y \le 1) = .594 - .264 = .330$

14.

a. $P(X_1 < t, X_2 < t, \ldots, X_{10} < t) = P(X_1 < t) \ldots P(X_{10} < t) = (1-e^{-lt})^{10}$

b. If "success" = {fail before t}, then p = P(success) = $1-e^{-lt}$,

and P(k successes among 10 trials) = $\binom{10}{k}1-e^{-lt^{k}}(e^{-lt})^{10-k}$

c. P(exactly 5 fail) = P(5 of l's fail and other 5 don't) + P(4 of l's fail, m fails, and other 5 don't) = $\binom{9}{5}(1-e^{-lt})^5(e^{-lt})^4(e^{-mt}) + \binom{9}{4}(1-e^{-lt})^4(1-e^{-mt})(e^{-lt})^5$

15.

a. F(y) = P(Y ≤ y) = P [(X$_1$ ≤y) ∪ ((X$_2$ ≤ y) ∩ (X$_3$ ≤ y))]
= P (X$_1$ ≤ y) + P[(X$_2$ ≤ y) ∩ (X$_3$ ≤ y)] - P[(X$_1$ ≤ y) ∩ (X$_2$ ≤ y) ∩ (X$_3$ ≤ y)]
= $(1-e^{-ly})+(1-e^{-ly})^2-(1-e^{-ly})^3$ for y ≥ 0

b. f(y) = F′(y) = $le^{-ly}+2(1-e^{-ly})\left(le^{-ly}\right)-3(1-e^{-ly})^2\left(le^{-ly}\right)$
= $4le^{-2ly}-3le^{-3ly}$ for y ≥ 0

E(Y) = $\int_0^\infty y\cdot\left(4le^{-2ly}-3le^{-3ly}\right)dy=2\left(\frac{1}{2l}\right)-\frac{1}{3l}=\frac{2}{3l}$

16.

a. f(x$_1$, x$_3$) = $\int_{-\infty}^{\infty} f(x_1,x_2,x_3)dx_2=\int_0^{1-x_1-x_3} kx_1x_2(1-x_3)dx_2$
$72x_1(1-x_3)(1-x_1-x_3)^2$ 0 ≤ x$_1$, 0 ≤ x$_3$, x$_1$ + x$_3$ ≤ 1

b. P(X$_1$ + X$_3$ ≤ .5) = $\int_0^{.5}\int_0^{.5-x_1} 72x_1(1-x_3)(1-x_1-x_3)^2\,dx_2dx_1$
= (after much algebra) .53125

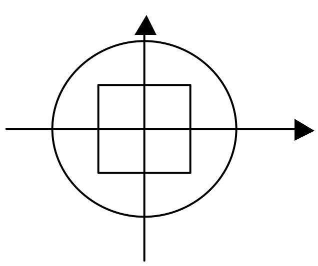
c. $f_{x_1}(x_1)=\int_{-\infty}^{\infty} f(x_1,x_3)dx_3=\int 72x_1(1-x_3)(1-x_1-x_3)^2\,dx_3$
$18x_1-48x_1^2+36x_1^3-6x_1^5$ 0 ≤ x$_1$ ≤ 1

17.

a. $P((X,Y)$ within a circle of radius $\frac{R}{2})=P(A)=\iint_A f(x,y)dxdy$

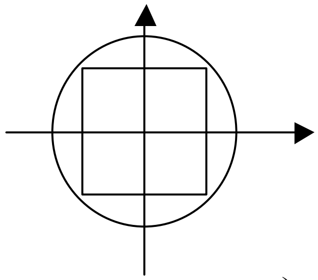
$$=\frac{1}{pR^2}\iint_A dxdy=\frac{area.of.A}{pR^2}=\frac{1}{4}=.25$$

b.



$$P\left(-\frac{R}{2}\le X\le\frac{R}{2},-\frac{R}{2}\le Y\le\frac{R}{2}\right)=\frac{R^2}{pR^2}=\frac{1}{p}$$

c.



$$P\left(-\frac{R}{\sqrt{2}} \le X \le \frac{R}{\sqrt{2}}, -\frac{R}{\sqrt{2}} \le Y \le \frac{R}{\sqrt{2}}\right) = \frac{2R^2}{pR^2} = \frac{2}{p}$$

d. $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{pR^2}dy = \frac{2\sqrt{R^2-x^2}}{pR^2}$ for $-R \le x \le R$ and similarly for $f_Y(y)$. X and Y are not independent since e.g. $f_x(.9R) = f_Y(.9R) > 0$, yet $f(.9R, .9R) = 0$ since (.9R, .9R) is outside the circle of radius R.

18.

a. $P_{y|X}(y|1)$ results from dividing each entry in x = 1 row of the joint probability table by $p_x(1) = .34$:

$$P_{y|x}(0\,|\,1) = \frac{.08}{.34} = .2353$$

$$P_{y|x}(1\,|\,1) = \frac{.20}{.34} = .5882$$

$$P_{y|x}(2\,|\,1) = \frac{.06}{.34} = .1765$$

b. $P_{y|X}(x|2)$ is requested; to obtain this divide each entry in the y = 2 row by $p_x(2) = .50$:

y	0	1	2
$P_{y\|X}(y\|2)$	.12	.28	.60

c. $P(Y \le 1 \mid x = 2) = P_{y|X}(0|2) + P_{y|X}(1|2) = .12 + .28 = .40$

d. $P_{X|Y}(x|2)$ results from dividing each entry in the y = 2 column by $p_y(2) = .38$:

x	0	1	2
$P_{x\|y}(x\|2)$	.0526	.1579	.7895

19.

a. $f_{Y|X}(y \mid x) = \dfrac{f(x,y)}{f_X(x)} = \dfrac{k(x^2+y^2)}{10kx^2+.05}$ $\quad 20 \le y \le 30$

$f_{X|Y}(x \mid y) = \dfrac{k(x^2+y^2)}{10ky^2+.05}$ $\quad 20 \le x \le 30$ $\quad \left(k = \dfrac{3}{380{,}000}\right)$

b. P(Y ≥ 25 | X = 22) = $\int_{25}^{30} f_{Y|X}(y \mid 22)\,dy$

$= \int_{25}^{30} \dfrac{k((22)^2+y^2)}{10k(22)^2+.05}\,dy = .783$

P(Y ≥ 25) = $\int_{25}^{30} f_Y(y)\,dy = \int_{25}^{30} (10ky^2+.05)\,dy = .75$

c. E(Y | X=22) = $\int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid 22)\,dy = \int_{20}^{30} y \cdot \dfrac{k((22)^2+y^2)}{10k(22)^2+.05}\,dy$

= 25.372912

E(Y^2 | X=22) = $\int_{20}^{30} y^2 \cdot \dfrac{k((22)^2+y^2)}{10k(22)^2+.05}\,dy = 652.028640$

V(Y| X = 22) = E(Y^2 | X=22) – [E(Y | X=22)]^2 = 8.243976

20.

a. $f_{x_3|x_1,x_2}(x_3 \mid x_1, x_2) = \dfrac{f(x_1,x_2,x_3)}{f_{x_1,x_2}(x_1,x_2)}$ where $f_{x_1,x_2}(x_1,x_2) =$ the marginal joint pdf of (X_1, X_2) = $\int_{-\infty}^{\infty} f(x_1,x_2,x_3)\,dx_3$

b. $f_{x_2,x_3|x_1}(x_2,x_3 \mid x_1) = \dfrac{f(x_1,x_2,x_3)}{f_{x_1}(x_1)}$ where

$f_{x_1}(x_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2,x_3)\,dx_2\,dx_3$

21. For every x and y, $f_{Y|X}(y|x) = f_y(y)$, since then $f(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = f_Y(y) \cdot f_X(x)$, as required.

Section 5.2

22.

a. E(X + Y) = $\sum_x \sum_y (x+y)p(x,y) = (0+0)(.02)$

$+(0+5)(.06)+...+(10+15)(.01) = 14.10$

b. E[max (X,Y)] = $\sum_x \sum_y \max(x+y)\cdot p(x,y)$

$= (0)(.02)+(5)(.06)+...+(15)(.01) = 9.60$

23. E(X_1 – X_2) = $\sum_{x_1=0}^{4}\sum_{x_2=0}^{3}(x_1-x_2)\cdot p(x_1,x_2) =$

(0 – 0)(.08) + (0 – 1)(.07) + … + (4 – 3)(.06) = .15
(which also equals E(X_1) – E(X_2) = 1.70 – 1.55)

24. Let h(X,Y) = # of individuals who handle the message.

		y					
	h(x,y)	1	2	3	4	5	6
	1	-	2	3	4	3	2
	2	2	-	2	3	4	3
x	3	3	2	-	2	3	4
	4	4	3	2	-	2	3
	5	3	4	3	2	-	2
	6	2	3	4	3	2	-

Since p(x,y) = $\frac{1}{30}$ for each possible (x,y), E[h(X,Y)] = $\sum_x \sum_y h(x,y)\cdot \frac{1}{30} = \frac{84}{30} = 2.80$

25. E(XY) = E(X) · E(Y) = L · L = L^2

26. Revenue = 3X + 10Y, so E (revenue) = E (3X + 10Y)

$$= \sum_{x=0}^{5}\sum_{y=0}^{2}(3x+10y)\cdot p(x,y) = 0\cdot p(0,0)+...+35\cdot p(5,2) = 15.4$$

27. E[h(X,Y)] = $\int_0^1\int_0^1 |x-y|\cdot 6x^2ydxdy = 2\int_0^1\int_0^x (x-y)\cdot 6x^2ydydx$

$$12\int_0^1\int_0^x \left(x^3y - x^2y^2\right)dydx = 12\int_0^1 \frac{x^5}{6}dx = \frac{1}{3}$$

28. E(XY) = $\sum_x\sum_y xy\cdot p(x,y) = \sum_x\sum_y xy\cdot p_x(x)\cdot p_y(y) = \sum_x xp_x(x)\cdot\sum_y yp_y(y)$

= E(X) · E(Y). (replace Σ with $\int$ in the continuous case)

29. Cov(X,Y) = $-\frac{2}{75}$ and $m_x = m_y = \frac{2}{5}$. E(X^2) = $\int_0^1 x^2\cdot f_x(x)dx$

$= 12\int_0^1 x^3(1-x^2dx) = \frac{12}{60} = \frac{1}{5}$, so Var (X) = $\frac{1}{5} - \frac{4}{25} = \frac{1}{25}$

Similarly, Var(Y) = $\frac{1}{25}$, so $r_{X,Y} = \frac{-2/75}{\sqrt{\frac{1}{25}}\cdot\sqrt{\frac{1}{25}}} = -\frac{50}{75} = -.667$

30.

a. E(X) = 5.55, E(Y) = 8.55, E(XY) = (0)(.02) + (0)(.06) + … + (150)(.01) = 44.25, so Cov(X,Y) = 44.25 – (5.55)(8.55) = -3.20

b. $s_X^2 = 12.45, s_Y^2 = 19.15$, so $r_{X,Y} = \frac{-3.20}{\sqrt{(12.45)(19.15)}} = -.207$

31.

a. E(X) = $\int_{20}^{30} xf_x(x)dx = \int_{20}^{30} x\left[10Kx^2 + .05\right]dx = 25.329 = E(Y)$

E(XY) = $\int_{20}^{30}\int_{20}^{30} xy\cdot K(x^2+y^2)dxdy = 641.447$

$\Rightarrow Cov(X,Y) = 641.447 - (25.329)^2 = -.111$

b. E(X^2) = $\int_{20}^{30} x^2\left[10Kx^2 + .05\right]dx = 649.8246 = E(Y^2)$,

so Var (X) = Var(Y) = 649.8246 – $(25.329)^2$ = 8.2664

$\Rightarrow r = \frac{-.111}{\sqrt{(8.2664)(8.2664)}} = -.0134$

32. There is a difficulty here. Existence of r requires that both X and Y have finite means and variances. Yet since the marginal pdf of Y is $\frac{1}{(1-y)^2}$ for y ≥ 0,

$$E(y)=\int_0^\infty \frac{y}{(1+y)^2}dy=\int_0^\infty \frac{(1+y-1)}{(1+y)^2}dy=\int_0^\infty \frac{1}{(1+y)}dy-\int_0^\infty \frac{1}{(1+y)^2}dy$$, and the first integral is not finite. Thus r itself is undefined.

33. Since E(XY) = E(X) · E(Y), Cov(X,Y) = E(XY) – E(X) · E(Y) = E(X) · E(Y) - E(X) · E(Y) = 0, and since Corr(X,Y) = $\frac{Cov(X,Y)}{s_x s_y}$, then Corr(X,Y) = 0

34.

a. In the discrete case, Var[h(X,Y)] = E{[h(X,Y) – E(h(X,Y))]^2} =

$$\sum_x\sum_y [h(x,y)-E(h(X,Y))]^2 p(x,y)=\sum_x\sum_y [h(x,y)^2 p(x,y)]-[E(h(X,Y))]^2$$

with $\iint$ replacing $\sum\sum$ in the continuous case.

b. E[h(X,Y)] = E[max(X,Y)] = 9.60, and E[h^2(X,Y)] = E[(max(X,Y))^2] = $(0)^2(.02)$ $+(5)^2(.06) + \ldots + (15)^2(.01) = 105.5$, so Var[max(X,Y)] = 105.5 – $(9.60)^2$ = 13.34

35.

a. Cov(aX + b, cY + d) = E[(aX + b)(cY + d)] – E(aX + b) · E(cY + d)
= E[acXY + adX + bcY + bd] – (aE(X) + b)(cE(Y) + d)
= acE(XY) – acE(X)E(Y) = acCov(X,Y)

b. Corr(aX + b, cY + d) =

$$\frac{Cov(aX+b,cY+d)}{\sqrt{Var(aX+b)}\sqrt{Var(cY+d)}}=\frac{acCov(X,Y)}{|a|\cdot|c|\sqrt{Var(X)\cdot Var(Y)}}$$

= Corr(X,Y) when a and c have the same signs.

c. When a and c differ in sign, Corr(aX + b, cY + d) = -Corr(X,Y).

36. Cov(X,Y) = Cov(X, aX+b) = E[X·(aX+b)] – E(X) ·E(aX+b) = a Var(X),

so Corr(X,Y) = $\frac{aVar(X)}{\sqrt{Var(X)\cdot Var(Y)}}=\frac{aVar(X)}{\sqrt{Var(X)\cdot a^2Var(X)}}$ = 1 if a > 0, and –1 if a < 0

Section 5.3

37.

	$P(x_1)$	.20	.50	.30
$P(x_2)$	$x_2 \mid x_1$	25	40	65
.20	25	.04	.10	.06
.50	40	.10	.25	.15
.30	65	.06	.15	.09

a.

$\bar{x}$	25	32.5	40	45	52.5	65
$p(\bar{x})$	.04	.20	.25	.12	.30	.09

$$E(\bar{x}) = (25)(.04) + 32.5(.20) + \ldots + 65(.09) = 44.5 = m$$

b.

s^2	0	112.5	312.5	800
$P(s^2)$	.38	.20	.30	.12

$E(s^2) = 212.25 = \sigma^2$

38.

a.

T_0	0	1	2	3	4
$P(T_0)$	.04	.20	.37	.30	.09

b. $m_{T_0} = E(T_0) = 2.2 = 2 \cdot m$

c. $s^2_{T_0} = E(T_0^2) - E(T_0)^2 = 5.82 - (2.2)^2 = .98 = 2 \cdot s^2$

39.

x	0	1	2	3	4	5	6	7	8	9	10
x/n	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
p(x/n)	.000	.000	.000	.001	.005	.027	.088	.201	.302	.269	.107

X is a binomial random variable with p = .8.

40.

a. Possible values of M are: 0, 5, 10. M = 0 when all 3 envelopes contain 0 money, hence $p(M = 0) = (.5)^3 = .125$. M = 10 when there is a single envelope with \$10, hence p(M = 10) = 1 – p(no envelopes with \$10) = $1 - (.8)^3 = .488$.
p(M = 5) = 1 – [.125 + .488] = .387.

M	0	5	10
p(M)	.125	.387	.488

An alternative solution would be to list all 27 possible combinations using a tree diagram and computing probabilities directly from the tree.

b. The statistic of interest is M, the maximum of x_1, x_2, or x_3, so that M = 0, 5, or 10. The population distribution is a s follows:

x	0	5	10
p(x)	1/2	3/10	1/5

Write a computer program to generate the digits 0 – 9 from a uniform distribution. Assign a value of 0 to the digits 0 – 4, a value of 5 to digits 5 – 7, and a value of 10 to digits 8 and 9. Generate samples of increasing sizes, keeping the number of replications constant and compute M from each sample. As n, the sample size, increases, p(M = 0) goes to zero, p(M = 10) goes to one. Furthermore, p(M = 5) goes to zero, but at a slower rate than p(M = 0).

41.

Outcome	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
Probability	.16	.12	.08	.04	.12	.09	.06	.03
$\overline{x}$	1	1.5	2	2.5	1.5	2	2.5	3
r	0	1	2	3	1	0	1	2

Outcome	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4
Probability	.08	.06	.04	.02	.04	.03	.02	.01
$\overline{x}$	2	2.5	3	3.5	2.5	3	3.5	4
r	2	1	0	1	3	2	1	2

a.

$\overline{x}$	1	1.5	2	2.5	3	3.5	4
$p(\overline{x})$	.16	.24	.25	.20	.10	.04	.01

b. $P(\overline{x} \le 2.5) = .8$

c.

r	0	1	2	3
p(r)	.30	.40	.22	.08

d. $P(\overline{X} \le 1.5)$ = P(1,1,1,1) + P(2,1,1,1) + … + P(1,1,1,2) + P(1,1,2,2) + … + P(2,2,1,1) + P(3,1,1,1) + … + P(1,1,1,3)
$= (.4)^4 + 4(.4)^3(.3) + 6(.4)^2(.3)^2 + 4(.4)^2(.2)^2 = .2400$

42.

a.

$\overline{x}$	27.75	28.0	29.7	29.95	31.65	31.9	33.6
$p(\overline{x})$	$\frac{4}{30}$	$\frac{2}{30}$	$\frac{6}{30}$	$\frac{4}{30}$	$\frac{8}{30}$	$\frac{4}{30}$	$\frac{2}{30}$

b.

$\overline{x}$	27.75	31.65	31.9
$p(\overline{x})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

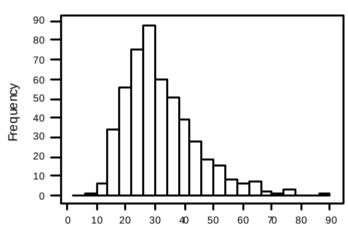
c. all three values are the same: 30.4333

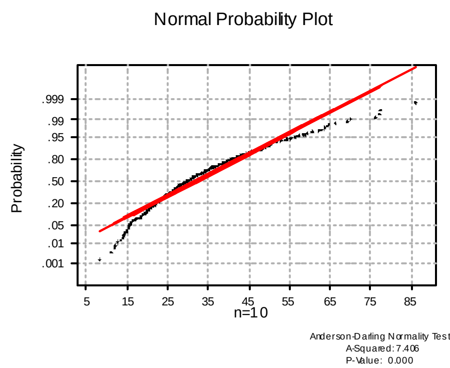
43. The statistic of interest is the fourth spread, or the difference between the medians of the upper and lower halves of the data. The population distribution is uniform with A = 8 and B = 10. Use a computer to generate samples of sizes n = 5, 10, 20, and 30 from a uniform distribution with A = 8 and B = 10. Keep the number of replications the same (say 500, for example). For each sample, compute the upper and lower fourth, then compute the difference. Plot the sampling distributions on separate histograms for n = 5, 10, 20, and 30.

44. Use a computer to generate samples of sizes n = 5, 10, 20, and 30 from a Weibull distribution with parameters as given, keeping the number of replications the same, as in problem 43 above. For each sample, calculate the mean. Below is a histogram, and a normal probability plot for the sampling distribution of $\overline{x}$ for n = 5, both generated by Minitab. This sampling distribution appears to be normal, so since larger sample sizes will produce distributions that are closer to normal, the others will also appear normal.

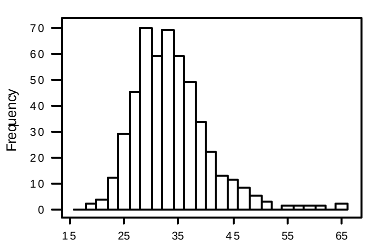
45. Using Minitab to generate the necessary sampling distribution, we can see that as n increases, the distribution slowly moves toward normality. However, even the sampling distribution for n = 50 is not yet approximately normal.

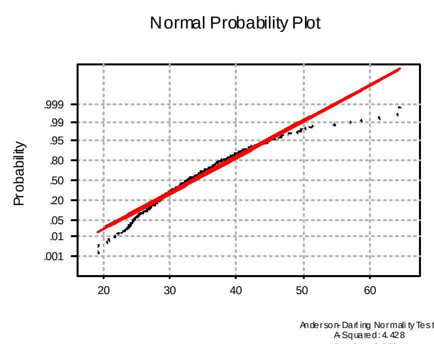
n = 10





n = 50





Section 5.4

46. μ = 12 cm σ = .04 cm

a. n = 16 $E(\overline{X}) = m = 12cm \quad s_{\overline{x}} = \frac{s_x}{\sqrt{n}} = \frac{.04}{4} = .01cm$

b. n = 64 $E(\overline{X}) = m = 12cm \quad s_{\overline{x}} = \frac{s_x}{\sqrt{n}} = \frac{.04}{8} = .005cm$

c. $\overline{X}$ is more likely to be within .01 cm of the mean (12 cm) with the second, larger, sample. This is due to the decreased variability of $\overline{X}$ with a larger sample size.

47. μ = 12 cm σ = .04 cm

a. n = 16 P(11.99 ≤ $\overline{X}$ ≤ 12.01) = $P\left(\frac{11.99-12}{.01} \le Z \le \frac{12.01-12}{.01}\right)$

= P(-1 ≤ Z ≤ 1)
= Φ(1) - Φ(-1)
=.8413 - .1587
=.6826

b. n = 25 P($\overline{X}$ > 12.01) = $P\left(Z > \frac{12.01-12}{.04/5}\right)$ = P(Z > 1.25)

= 1 - Φ(1.25)
= 1 - .8944
=.1056

48.

a. $m_{\overline{X}} = m = 50$, $s_{\overline{x}} = \frac{s_x}{\sqrt{n}} = \frac{1}{\sqrt{100}} = .10$

P(49.75 ≤ $\overline{X}$ ≤ 50.25) = $P\left(\frac{49.75-50}{.10} \le Z \le \frac{50.25-50}{.10}\right)$

= P(-2.5 ≤ Z ≤ 2.5) = .9876

b. P(49.75 ≤ $\overline{X}$ ≤ 50.25) ≈ $P\left(\frac{49.75-49.8}{.10} \le Z \le \frac{50.25-49.8}{.10}\right)$

= P(-.5 ≤ Z ≤ 4.5) = .6915

49.

a. 11 P.M. – 6:50 P.M. = 250 minutes. With $T_0 = X_1 + \dots + X_{40}$ = total grading time,

$m_{T_0} = nm = (40)(6) = 240$ and $s_{T_0} = s\sqrt{n} = 37.95,$ so $P(T_0 \le 250) \approx$

$$P\left(Z \le \frac{250-240}{37.95}\right) = P(Z \le .26) = .6026$$

b. $$P(T_0 > 260) = P\left(Z > \frac{260-240}{37.95}\right) = P(Z > .53) = .2981$$

50. $\mu = 10{,}000$ psi $\sigma = 500$ psi

a. n = 40

$P(9{,}900 \le \overline{X} \le 10{,}200) \approx P\left(\frac{9{,}900-10{,}000}{500/\sqrt{40}} \le Z \le \frac{10{,}200-10{,}000}{500/\sqrt{40}}\right)$

$= P(-1.26 \le Z \le 2.53)$

$= \Phi(2.53) - \Phi(-1.26)$

$= .9943 - .1038$

$= .8905$

b. According to the Rule of Thumb given in Section 5.4, n should be greater than 30 in order to apply the C.L.T., thus using the same procedure for n = 15 as was used for n = 40 would not be appropriate.

51. $X \sim N(10,4)$. For day 1, n = 5

$$P(\overline{X} \le 11) = P\left(Z \le \frac{11-10}{2/\sqrt{5}}\right) = P(Z \le 1.12) = .8686$$

For day 2, n = 6

$$P(\overline{X} \le 11) = P\left(Z \le \frac{11-10}{2/\sqrt{6}}\right) = P(Z \le 1.22) = .8888$$

For both days,

$P(\overline{X} \le 11) = (.8686)(.8888) = .7720$

52. $X \sim N(10)$, n =4

$m_{T_0} = nm = (4)(10) = 40$ and $s_{T_0} = s\sqrt{n} = (2)(1) = 2,$

We desire the 95th percentile: $40 + (1.645)(2) = 43.29$

53. μ = 50, σ = 1.2

a. n = 9

P($\overline{X} \geq 51) = P\left(Z \geq \dfrac{51-50}{1.2/\sqrt{9}}\right) = P(Z \geq 2.5) = 1 - .9938 = .0062$

b. n = 40

P($\overline{X} \geq 51) = P\left(Z \geq \dfrac{51-50}{1.2/\sqrt{40}}\right) = P(Z \geq 5.27) \approx 0$

54.

a. $m_{\overline{X}} = m = 2.65$, $s_{\bar{x}} = \dfrac{s_x}{\sqrt{n}} = \dfrac{.85}{5} = .17$

P($\overline{X} \leq 3.00) = P\left(Z \leq \dfrac{3.00-2.65}{.17}\right) = P(Z \leq 2.06) = .9803$

P(2.65 ≤ $\overline{X} \leq 3.00) = = P(\overline{X} \leq 3.00) - P(\overline{X} \leq 2.65) = .4803$

b. P($\overline{X} \leq 3.00) = P\left(Z \leq \dfrac{3.00-2.65}{.85/\sqrt{n}}\right) = .99$ implies that $\dfrac{.35}{85/\sqrt{n}} = 2.33,$ from which n = 32.02. Thus n = 33 will suffice.

55. $m = np = 20 \quad s = \sqrt{npq} = 3.464$

a. P(25 ≤ X) $\approx P\left(\dfrac{24.5-20}{3.464} \leq Z\right) = P(1.30 \leq Z) = .0968$

b. P(15 ≤ X ≤ 25) $\approx P\left(\dfrac{14.5-20}{3.464} \leq Z \leq \dfrac{25.5-20}{3.464}\right)$

$= P(-1.59 \leq Z \leq 1.59) = .8882$

56.

a. With Y = # of tickets, Y has approximately a normal distribution with $m = l = 50$, $s = \sqrt{l} = 7.071$, so P(35 ≤ Y ≤ 70) $\approx P\left(\dfrac{34.5-50}{7.071} \leq Z \leq \dfrac{70.5-50}{7.071}\right)$ = P(-2.19 ≤ Z ≤ 2.90) = .9838

b. Here $m = 250$, $s^2 = 250, s = 15.811$, so P(225 ≤ Y ≤ 275) ≈ $P\left(\dfrac{224.5-250}{15.811} \leq Z \leq \dfrac{275.5-250}{15.811}\right)$ = P(-1.61 ≤ Z ≤ 1.61) = .8926

57. E(X) = 100, Var(X) = 200, $s_x = 14.14$, so P(X ≤ 125) ≈ $P\left(Z \le \dfrac{125-100}{14.14}\right)$

= P(Z ≤ 1.77) = .9616

Section 5.5

58.

a. E($27X_1 + 125X_2 + 512X_3$) = 27 E(X_1) + 125 E(X_2) + 512 E(X_3)
= 27(200) + 125(250) + 512(100) = 87,850
V($27X_1 + 125X_2 + 512X_3$) = 27^2 V(X_1) + 125^2 V(X_2) + 512^2 V(X_3)
= 27^2 $(10)^2$ + 125^2 $(12)^2$ + 512^2 $(8)^2$ = 19,100,116

b. The expected value is still correct, but the variance is not because the covariances now also contribute to the variance.

59.

a. E($X_1 + X_2 + X_3$) = 180, V($X_1 + X_2 + X_3$) = 45, $s_{x_1+x_2+x_3} = 6.708$

P($X_1 + X_2 + X_3 \le 200$) = $P\left(Z \le \dfrac{200-180}{6.708}\right) = P(Z \le 2.98) = .9986$

P($150 \le X_1 + X_2 + X_3 \le 200$) = $P(-4.47 \le Z \le 2.98) \approx .9986$

b. $m_{\overline{X}} = m = 60$, $s_{\bar{x}} = \dfrac{s_x}{\sqrt{n}} = \dfrac{\sqrt{15}}{\sqrt{3}} = 2.236$

$$P(\overline{X} \ge 55) = P\left(Z \ge \frac{55-60}{2.236}\right) = P(Z \ge -2.236) = .9875$$

$$P(58 \le \overline{X} \le 62) = P(-.89 \le Z \le .89) = .6266$$

c. E(X_1 - $.5X_2$ -$.5X_3$) = 0;

V(X_1 - $.5X_2$ -$.5X_3$) = $s_1^2 + .25s_2^2 + .25s_3^2 = 22.5$, sd = 4.7434

P($-10 \le X_1 - .5X_2 - .5X_3 \le 5$) = $P\left(\dfrac{-10-0}{4.7434} \le Z \le \dfrac{5-0}{4.7434}\right)$

$= P(-2.11 \le Z \le 1.05)$ = .8531 - .0174 = .8357

d. E($X_1 + X_2 + X_3$) = 150, V($X_1 + X_2 + X_3$) = 36, $\sigma_{x_1+x_2+x_3} = 6$

P($X_1 + X_2 + X_3 \le 200$) = $P\left(Z \le \dfrac{160-150}{6}\right) = P(Z \le 1.67) = .9525$

We want P($X_1 + X_2 \ge 2X_3$), or written another way, P($X_1 + X_2 - 2X_3 \ge 0$)
E($X_1 + X_2 - 2X_3$) = 40 + 50 – 2(60) = -30,
V($X_1 + X_2 - 2X_3$) = $\sigma_1^2 + \sigma_2^2 + 4\sigma_3^2 = 78$, 36, sd = 8.832, so

P($X_1 + X_2 - 2X_3 \ge 0$) = $P\left(Z \ge \dfrac{0-(-30)}{8.832}\right) = P(Z \ge 3.40) = .0003$

60. Y is normally distributed with $\mu_Y = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5) = -1$, and

$\sigma_Y^2 = \frac{1}{4}\sigma_1^2 + \frac{1}{4}\sigma_2^2 + \frac{1}{9}\sigma_3^2 + \frac{1}{9}\sigma_4^2 + \frac{1}{9}\sigma_5^2 = 3.167, \sigma_Y = 1.7795$.

Thus, $P(0 \le Y) = P\left(\dfrac{0-(-1)}{1.7795} \le Z\right) = P(.56 \le Z) = .2877$ and

$P(-1 \le Y \le 1) = P\left(0 \le Z \le \dfrac{2}{1.7795}\right) = P(0 \le Z \le 1.12) = .3686$

61.

a. The marginal pmf's of X and Y are given in the solution to Exercise 7, from which E(X) = 2.8, E(Y) = .7, V(X) = 1.66, V(Y) = .61. Thus E(X+Y) = E(X) + E(Y) = 3.5, V(X+Y) = V(X) + V(Y) = 2.27, and the standard deviation of X + Y is 1.51

b. E(3X+10Y) = 3E(X) + 10E(Y) = 15.4, V(3X+10Y) = 9V(X) + 100V(Y) = 75.94, and the standard deviation of revenue is 8.71

62. E($X_1 + X_2 + X_3$) = E(X_1) + E(X_2) + E(X_3) = 15 + 30 + 20 = 65 min.,
V($X_1 + X_2 + X_3$) = $1^2 + 2^2 + 1.5^2$ = 7.25, $\sigma_{x_1+x_2+x_3} = \sqrt{7.25} = 2.6926$

Thus, P($X_1 + X_2 + X_3 \le 60$) = $P\left(Z \le \dfrac{60-65}{2.6926}\right) = P(Z \le -1.86) = .0314$

63.

a. E(X_1) = 1.70, E(X_2) = 1.55, E(X_1X_2) = $\sum_{x_1}\sum_{x_2} x_1 x_2 p(x_1, x_2) = 3.33$, so Cov($X_1$,$X_2$) = E($X_1X_2$) - E($X_1$) E($X_2$) = 3.33 – 2.635 = .695

b. V($X_1 + X_2$) = V(X_1) + V(X_2) + 2 Cov(X_1,X_2)
= 1.59 + 1.0875 + 2(.695) = 4.0675

64. Let $X_1, \ldots, X_5$ denote morning times and $X_6, \ldots, X_{10}$ denote evening times.

a. $E(X_1 + \ldots + X_{10}) = E(X_1) + \ldots + E(X_{10}) = 5\,E(X_1) + 5\,E(X_6)$
$= 5(4) + 5(5) = 45$

b. $Var(X_1 + \ldots + X_{10}) = Var(X_1) + \ldots + Var(X_{10}) = 5\,Var(X_1) + 5Var(X_6)$

$$= 5\left[\frac{64}{12} + \frac{100}{12}\right] = \frac{820}{12} = 68.33$$

c. $E(X_1 - X_6) = E(X_1) - E(X_6) = 4 - 5 = -1$

$Var(X_1 - X_6) = Var(X_1) + Var(X_6) = \dfrac{64}{12} + \dfrac{100}{12} = \dfrac{164}{12} = 13.67$

d. $E[(X_1 + \ldots + X_5) - (X_6 + \ldots + X_{10})] = 5(4) - 5(5) = -5$
$Var[(X_1 + \ldots + X_5) - (X_6 + \ldots + X_{10})]$
$= Var(X_1 + \ldots + X_5) + Var(X_6 + \ldots + X_{10})] = 68.33$

65. $\mu = 5.00$, $\sigma = .2$

a. $E(\overline{X} - \overline{Y}) = 0; \qquad V(\overline{X} - \overline{Y}) = \dfrac{s^2}{25} + \dfrac{s^2}{25} = .0032,\ s_{\overline{X}-\overline{Y}} = .0566$

$$\Rightarrow P\left(-.1 \le \overline{X} - \overline{Y} \le .1\right) \approx P\left(-1.77 \le Z \le 1.77\right) = .9232 \text{ (by the CLT)}$$

b. $V(\overline{X} - \overline{Y}) = \dfrac{s^2}{36} + \dfrac{s^2}{36} = .0022222,\ s_{\overline{X}-\overline{Y}} = .0471$

$$\Rightarrow P\left(-.1 \le \overline{X} - \overline{Y} \le .1\right) \approx P\left(-2.12 \le Z \le 2.12\right) = .9660$$

66.

a. With $M = 5X_1 + 10X_2$, $E(M) = 5(2) + 10(4) = 50$,
$Var(M) = 5^2\,(.5)^2 + 10^2\,(1)^2 = 106.25$, $\sigma_M = 10.308$.

b. $P(\,75 < M\,) = P\left(\dfrac{75-50}{10.308} < Z\right) = P(2.43 < Z) = .0075$

c. $M = A_1X_1 + A_2X_2$ with the A_I's and X_I's all independent, so
$E(M) = E(A_1X_1) + E(A_2X_2) = E(A_1)E(X_1) + E(A_2)E(X_2) = 50$

d. $Var(M) = E(M^2) - [E(M)]^2$. Recall that for any r.v. Y,
$E(Y^2) = Var(Y) + [E(Y)]^2$. Thus, $E(M^2) = E\left(A_1^2X_1^2 + 2A_1X_1A_2X_2 + A_2^2X_2^2\right)$
$= E\left(A_1^2\right)E\left(X_1^2\right) + 2E\left(A_1\right)E\left(X_1\right)E\left(A_2\right)E\left(X_2\right) + E\left(A_2^2\right)E\left(X_2^2\right)$
(by independence)
$= (.25 + 25)(.25 + 4) + 2(5)(2)(10)(4) + (.25 + 100)(1 + 16) = 2611.5625$, so $Var(M) =$
$2611.5625 - (50)^2 = 111.5625$

e. E(M) = 50 still, but now
$$Var(M) = a_1^2 Var(X_1) + 2a_1 a_2 Cov(X_1, X_2) + a_2^2 Var(X_2)$$
= 6.25 + 2(5)(10)(-.25) + 100 = 81.25

67. Letting X_1, X_2, and X_3 denote the lengths of the three pieces, the total length is $X_1 + X_2 - X_3$. This has a normal distribution with mean value 20 + 15 – 1 = 34, variance .25+.16+.01 = .42, and standard deviation .6481. Standardizing gives
$P(34.5 \le X_1 + X_2 - X_3 \le 35) = P(.77 \le Z \le 1.54) = .1588$

68. Let X_1 and X_2 denote the (constant) speeds of the two planes.

a. After two hours, the planes have traveled $2X_1$ km. and $2X_2$ km., respectively, so the second will not have caught the first if $2X_1 + 10 > 2X_2$, i.e. if $X_2 - X_1 < 5$. $X_2 - X_1$ has a mean 500 – 520 = -20, variance 100 + 100 = 200, and standard deviation 14.14. Thus,
$$P(X_2 - X_1 < 5) = P\left(Z < \frac{5-(-20)}{14.14}\right) = P(Z < 1.77) = .9616.$$

b. After two hours, #1 will be $10 + 2X_1$ km from where #2 started, whereas #2 will be $2X_2$ from where it started. Thus the separation distance will be al most 10 if $|2X_2 - 10 - 2X_1| \le 10$, i.e. $-10 \le 2X_2 - 10 - 2X_1 \le 10$,
i.e. $0 \le X_2 - X_1 \le 10$. The corresponding probability is
$P(0 \le X_2 - X_1 \le 10) = P(1.41 \le Z \le 2.12) = .9830 - .9207 = .0623$.

69.

a. $E(X_1 + X_2 + X_3) = 800 + 1000 + 600 = 2400$.

b. Assuming independence of X_1, X_2 , X_3, $Var(X_1 + X_2 + X_3)$
$= (16)^2 + (25)^2 + (18)^2 = 12.05$

c. $E(X_1 + X_2 + X_3) =$ 2400 as before, but now $Var(X_1 + X_2 + X_3)$
$= Var(X_1) + Var(X_2) + Var(X_3) + 2Cov(X_1,X_2) + 2Cov(X_1, X_3) + 2Cov(X_2, X_3) = 1745$,
with sd = 41.77

70.

a. $E(Y_i) = .5,$ so $E(W) = \sum_{i=1}^{n} i \cdot E(Y_i) = .5\sum_{i=1}^{n} i = \frac{n(n+1)}{4}$

b. $Var(Y_i) = .25,$ so $Var(W) = \sum_{i=1}^{n} i^2 \cdot Var(Y_i) = .25\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{24}$

71.

a. $M = a_1X_1 + a_2X_2 + W\int_0^{12} x\,dx = a_1X_1 + a_2X_2 + 72W$, so

E(M) = (5)(2) + (10)(4) + (72)(1.5) = 158m

$s_M^2 = (5)^2(.5)^2 + (10)^2(1)^2 + (72)^2(.25)^2 = 430.25$, $s_M = 20.74$

b. $P(M \le 200) = P\left(Z \le \dfrac{200-158}{20.74}\right) = P(Z \le 2.03) = .9788$

72. The total elapsed time between leaving and returning is $T_o = X_1 + X_2 + X_3 + X_4$, with $E(T_o) = 40,\ s_{T_o}^2 = 40$, $s_{T_o} = 5.477$. T_o is normally distributed, and the desired value t is the 99th percentile of the lapsed time distribution added to 10 A.M.: 10:00 + [40+(5.477)(2.33)] = 10:52.76

73.

a. Both approximately normal by the C.L.T.

b. The difference of two r.v.'s is just a special linear combination, and a linear combination of normal r.v's has a normal distribution, so $\overline{X} - \overline{Y}$ has approximately a normal distribution with $m_{\overline{X}-\overline{Y}} = 5$ and $s_{\overline{X}-\overline{Y}}^2 = \dfrac{8^2}{40} + \dfrac{6^2}{35} = 2.629,\ s_{\overline{X}-\overline{Y}} = 1.621$

c. $P\left(-1 \le \overline{X} - \overline{Y} \le 1\right) \& P\left(\dfrac{-1-5}{1.6213} \le Z \le \dfrac{1-5}{1.6213}\right)$

$= P(-3.70 \le Z \le -2.47) \approx .0068$

d. $P\left(\overline{X} - \overline{Y} \ge 10\right) \& P\left(Z \ge \dfrac{10-5}{1.6213}\right) = P(Z \ge 3.08) = .0010.$ This probability is quite small, so such an occurrence is unlikely if $m_1 - m_2 = 5$, and we would thus doubt this claim.

74. X is approximately normal with $m_1 = (50)(.7) = 35$ and $s_1^2 = (50)(.7)(.3) = 10.5$, as is Y with $m_2 = 30$ and $s_2^2 = 12$. Thus $m_{X-Y} = 5$ and $s_{X-Y}^2 = 22.5$, so

$$p(-5 \le X - Y \le 5) \approx P\left(\frac{-10}{4.74} \le Z \le \frac{0}{4.74}\right) = P(-2.11 \le Z \le 0) = .4826$$

Supplementary Exercises

75.

a. $p_X(x)$ is obtained by adding joint probabilities across the row labeled x, resulting in $p_X(x)$ = .2, .5, .3 for x = 12, 15, 20 respectively. Similarly, from column sums $p_y(y)$ = .1, .35, .55 for y = 12, 15, 20 respectively.

b. $P(X \le 15 \text{ and } Y \le 15) = p(12,12) + p(12,15) + p(15,12) + p(15,15) = .25$

c. $p_x(12) \cdot p_y(12) = (.2)(.1) \neq .05 = p(12,12)$, so X and Y are not independent. (Almost any other (x,y) pair yields the same conclusion).

d. $E(X+Y) = \sum\sum (x+y)p(x,y) = 33.35$ (or = E(X) + E(Y) = 33.35)

e. $E(|X-Y|) = \sum\sum |x+y| p(x,y) = 3.85$

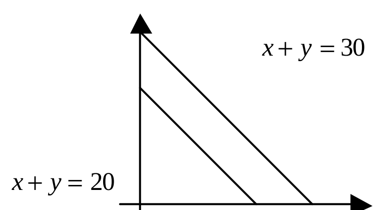
76. The roll-up procedure is not valid for the 75th percentile unless $s_1 = 0$ or $s_2 = 0$ or both s_1 and $s_2 = 0$, as described below.

Sum of percentiles: $m_1 + (Z)s_1 + m_2 + (Z)s_2 = m_1 + m_2 + (Z)(s_1 + s_2)$

Percentile of sums: $m_1 + m_2 + (Z)\sqrt{s_1^2 + s_2^2}$

These are equal when Z = 0 (i.e. for the median) or in the unusual case when $s_1 + s_2 = \sqrt{s_1^2 + s_2^2}$, which happens when $s_1 = 0$ or $s_2 = 0$ or both s_1 and $s_2 = 0$.

77.



a. $$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_0^{20}\int_{20-x}^{30-x} kxydydx + \int_{20}^{30}\int_0^{30-x} kxydydx$$
$$= \frac{81{,}250}{3} \cdot k \Rightarrow k = \frac{3}{81{,}250}$$

b. $$f_X(x) = \begin{cases} \int_{20-x}^{30-x} kxydy = k(250x - 10x^2) & 0 \le x \le 20 \\ \int_0^{30-x} kxydy = k(450x - 30x^2 + \tfrac{1}{2}x^3) & 20 \le x \le 30 \end{cases}$$

and by symmetry $f_Y(y)$ is obtained by substituting y for x in $f_X(x)$. Since $f_X(25) > 0$, and $f_Y(25) > 0$, but f(25, 25) = 0 , $f_X(x) \cdot f_Y(y) \neq f(x,y)$ for all x,y so X and Y are not independent.

c. $P(X+Y\le 25)=\int_0^{20}\int_{20-x}^{25-x} kxydydx+\int_{20}^{25}\int_0^{25-x} kxydydx$

$$=\frac{3}{81{,}250}\cdot\frac{230{,}625}{24}=.355$$

d. $E(X+Y)=E(X)+E(Y)=2\left\{\int_0^{20} x\cdot k\left(250x-10x^2\right)dx\right.$

$\left.+\int_{20}^{30} x\cdot k\left(450x-30x^2+\tfrac{1}{2}x^3\right)dx\right\} \quad =2k(351{,}666.67)=25.969$

e. $E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\cdot f(x,y)dxdy=\int_0^{20}\int_{20-x}^{30-x} kx^2y^2dydx$

$+\int_{20}^{30}\int_0^{30-x} kx^2y^2dydx=\frac{k}{3}\cdot\frac{33{,}250{,}000}{3}=136.4103$, so

Cov(X,Y) = 136.4103 – $(12.9845)^2$ = -32.19, and $E(X^2)=E(Y^2)$ = 204.6154, so

$s_x^2=s_y^2=204.6154-(12.9845)^2=36.0182$ and $r=\frac{-32.19}{36.0182}=-.894$

f. Var (X + Y) = Var(X) + Var(Y) + 2Cov(X,Y) = 7.66

78. $F_Y(y)=P(\max(X_1,\ldots,X_n)\le y)=P(X_1\le y,\ldots,X_n\le y)=[P(X_1\le y)]^n=\left(\frac{y-100}{100}\right)^n$ for $100\le y\le 200$.

Thus $f_Y(y)=\frac{n}{100^n}(y-100)^{n-1}$ for $100\le y\le 200$.

$$E(Y)=\int_{100}^{200} y\cdot\frac{n}{100^n}(y-100)^{n-1}dy=\frac{n}{100^n}\int_0^{100}(u+100)u^{n-1}du$$

$$=100+\frac{n}{100^n}\int_0^{100}u^n du=100+100\frac{n}{n+1}=\frac{2n+1}{n+1}\cdot 100$$

79. $E(\overline{X}+\overline{Y}+\overline{Z})=500+900+2000=3400$

$Var(\overline{X}+\overline{Y}+\overline{Z})=\frac{50^2}{365}+\frac{100^2}{365}+\frac{180^2}{365}=123.014$, and the std dev = 11.09.

$P(\overline{X}+\overline{Y}+\overline{Z}\le 3500)=P(Z\le 9.0)\approx 1$

80.

a. Let $X_1, \ldots, X_{12}$ denote the weights for the business-class passengers and $Y_1, \ldots, Y_{50}$ denote the tourist-class weights. Then T = total weight
$= X_1 + \ldots + X_{12} + Y_1 + \ldots + Y_{50} = X + Y$
$E(X) = 12E(X_1) = 12(30) = 360$; $V(X) = 12V(X_1) = 12(36) = 432$.
$E(Y) = 50E(Y_1) = 50(40) = 2000$; $V(Y) = 50V(Y_1) = 50(100) = 5000$.
Thus $E(T) = E(X) + E(Y) = 360 + 2000 = 2360$
And $V(T) = V(X) + V(Y) = 432 + 5000 = 5432$, std dev = 73.7021

b. $$P(T \le 2500) = P\left(Z \le \frac{2500 - 2360}{73.7021}\right) = P(Z \le 1.90) = .9713$$

81.

a. $E(N) \cdot \mu = (10)(40) = 400$ minutes

b. We expect 20 components to come in for repair during a 4 hour period, so $E(N) \cdot \mu = (20)(3.5) = 70$

82. X ~ Bin (200, .45) and Y ~ Bin (300, .6). Because both n's are large, both X and Y are approximately normal, so X + Y is approximately normal with mean (200)(.45) + (300)(.6) = 270, variance 200(.45)(.55) + 300(.6)(.4) = 121.40, and standard deviation 11.02. Thus, $P(X + Y \ge 250) = P\left(Z \ge \frac{249.5 - 270}{11.02}\right) = P(Z \ge -1.86) = .9686$

83. $0.95 = P(m - .02 \le \overline{X} \le m + .02) = P\left(\frac{-.02}{.01/\sqrt{n}} \le Z \le \frac{.02}{.01/\sqrt{n}}\right)$
$= P\left(-.2\sqrt{n} \le Z \le .2\sqrt{n}\right)$, but $P(-1.96 \le Z \le 1.96) = .95$ so
$.2\sqrt{n} = 1.96 \Rightarrow n = 97.$ The C.L.T.

84. I have 192 oz. The amount which I would consume if there were no limit is $T_o = X_1 + \ldots + X_{14}$ where each X_I is normally distributed with $\mu = 13$ and $\sigma = 2$. Thus T_o is normal with $m_{T_o} = 182$ and $s_{T_o} = 7.483$, so $P(T_o < 192) = P(Z < 1.34) = .9099$.

85. The expected value and standard deviation of volume are 87,850 and 4370.37, respectively, so
$$P(volume \le 100{,}000) = P\left(Z \le \frac{100{,}000 - 87{,}850}{4370.37}\right) = P(Z \le 2.78) = .9973$$

86. The student will not be late if $X_1 + X_3 \le X_2$, i.e. if $X_1 - X_2 + X_3 \le 0$. This linear combination has mean –2, variance 4.25, and standard deviation 2.06, so
$$P(X_1 - X_2 + X_3 \le 0) = P\left(Z \le \frac{0 - (-2)}{2.06}\right) = P(Z \le .97) = .8340$$

87.

a. $Var(aX+Y) = a^2 s_x^2 + 2aCov(X,Y) + s_y^2 = a^2 s_x^2 + 2a s_X s_Y r + s_y^2.$

Substituting $a = \frac{s_Y}{s_X}$ yields $s_Y^2 + 2s_Y^2 r + s_Y^2 = 2s_Y^2(1-r) \geq 0$, so $r \geq -1$

b. Same argument as in **a**

c. Suppose $r = 1$. Then $Var(aX - Y) = 2s_Y^2(1-r) = 0$, which implies that $aX - Y = k$ (a constant), so $aX - Y = aX - k$, which is of the form $aX + b$.

88. $E(X+Y-t)^2 = \int_0^1\int_0^1 (x+y-t)^2 \cdot f(x,y)dxdy.$ To find the minimizing value of t, take the derivative with respect to t and equate it to 0:

$0 = \int_0^1\int_0^1 2(x+y-t)(-1)f(x,y) = 0 \Rightarrow \int_0^1\int_0^1 tf(x,y)dxdy = t$

$= \int_0^1\int_0^1 (x+y)\cdot f(x,y)dxdy = E(X+Y)$, so the best prediction is the individual's expected score (= 1.167).

89.

a. With Y = $X_1 + X_2$,

$$F_Y(y) = \int_0^y \left\{ \int_0^{y-x_1} \frac{1}{2^{n_1/2}\Gamma(n_1/2)} \cdot \frac{1}{2^{n_{21}/2}\Gamma(n_2/2)} \cdot x_1^{\frac{n_1}{2}-1} x_2^{\frac{n_2}{2}-1} e^{-\frac{x_1+x_2}{2}} dx_2 \right\} dx_1 .$$

But the inner integral can be shown to be equal to

$\frac{1}{2^{(n_1+n_2)/2}\Gamma((n_1+n_2)/2)} y^{[(n_1+n_2)/2]-1} e^{-y/2}$, from which the result follows.

b. By **a,** $Z_1^2 + Z_2^2$ is chi-squared with $n = 2$, so $(Z_1^2 + Z_2^2) + Z_3^2$ is chi-squared with $n = 3$, etc, until $Z_1^2 + ... + Z_n^2$ 9s chi-squared with $n = n$

c. $\frac{X_i - m}{s}$ is standard normal, so $\left[\frac{X_i - m}{s}\right]^2$ is chi-squared with $n = 1$, so the sum is chi-squared with $n = n$.

90.

a. Cov(X, Y + Z) = E[X(Y + Z)] – E(X) · E(Y + Z)
= E(XY) + E(XZ) – E(X) · E(Y) – E(X) · E(Z)
= E(XY) – E(X) · E(Y) + E(XZ) – E(X) · E(Z)
= Cov(X,Y) + Cov(X,Z).

b. $Cov(X_1 + X_2, Y_1 + Y_2) = Cov(X_1, Y_1) + Cov(X_1, Y_2) + Cov(X_2, Y_1) + Cov(X_2, Y_2)$ (apply **a** twice) = 16.

91.

a. $V(X_1) = V(W + E_1) = s_W^2 + s_E^2 = V(W + E_2) = V(X_2)$ and

$$Cov(X_1, X_2) = Cov(W + E_1, W + E_2) = Cov(W,W) + Cov(W, E_2) +$$
$$Cov(E_1, W) + Cov(E_1, E_2) = Cov(W,W) = V(W) = s_w^2.$$

Thus, $r = \dfrac{s_W^2}{\sqrt{s_W^2 + s_E^2} \cdot \sqrt{s_W^2 + s_E^2}} = \dfrac{s_W^2}{s_W^2 + s_E^2}$

b. $r = \dfrac{1}{1 + .0001} = .9999$

92.

a. Cov(X,Y) = Cov(A+D, B+E)
= Cov(A,B) + Cov(D,B) + Cov(A,E) + Cov(D,E)= Cov(A,B). Thus

$$Corr(X,Y) = \frac{Cov(A,B)}{\sqrt{s_A^2 + s_D^2} \cdot \sqrt{s_B^2 + s_E^2}}$$

$$= \frac{Cov(A,B)}{s_A s_B} \cdot \frac{s_A}{\sqrt{s_A^2 + s_D^2}} \cdot \frac{s_B}{\sqrt{s_B^2 + s_E^2}}$$

The first factor in this expression is Corr(A,B), and (by the result of exercise 70**a**) the second and third factors are the square roots of $Corr(X_1, X_2)$ and $Corr(Y_1, Y_2)$, respectively. Clearly, measurement error reduces the correlation, since both square-root factors are between 0 and 1.

b. $\sqrt{.8100} \cdot \sqrt{.9025} = .855$. This is disturbing, because measurement error substantially reduces the correlation.

93. $E(Y) \approx h(m_1, m_2, m_3, m_4) = 120\left[\frac{1}{10} + \frac{1}{15} + \frac{1}{20}\right] = 26$

The partial derivatives of $h(m_1, m_2, m_3, m_4)$ with respect to x_1, x_2, x_3, and x_4 are $-\dfrac{x_4}{x_1^2}$, $-\dfrac{x_4}{x_2^2}$, $-\dfrac{x_4}{x_3^2}$, and $\dfrac{1}{x_1} + \dfrac{1}{x_2} + \dfrac{1}{x_3}$, respectively. Substituting $x_1 = 10$, $x_2 = 15$, $x_3 = 20$, and $x_4 = 120$ gives –1.2, -.5333, -.3000, and .2167, respectively, so $V(Y) = (1)(-1.2)^2 + (1)(-.5333)^2 + (1.5)(-.3000)^2 + (4.0)(.2167)^2 = 2.6783$, and the approximate sd of y is 1.64.

94. The four second order partials are $\dfrac{2x_4}{x_1^3}$, $\dfrac{2x_4}{x_2^3}$, $\dfrac{2x_4}{x_3^3}$, and 0 respectively. Substitution gives $E(Y) = 26 + .1200 + .0356 + .0338 = 26.1894$.