COMM 215: BUSINESS STATISTICS

SOLUTIONS TO PRACTICE PROBLEMS 1

Descriptive Statistics

1. a.

Leaf-unit = 0.01

0	26, 45, 52, 81, 85, 88, 91, 94
1	01, 22, 42, 91, 92, 94
2	25, 33, 51, 62, 66, 93, 95, 95
3	51, 60, 72
4	91
5	19
6	71
7	96
8	81

Leaf Unit = 0.10

0	24588899
1	024999
2	23566999
3	567
4	9
5	1
6	7
7	9
8	8

b. i. $\bar{X}$ = 2.689 Median = 2.29 Mode = 2.95

ii. Q1=(.25)(30) = 7.5= the 8th ordered value = 0.94
Q3=(.75)(30) = 22.5= the 23rd ordered value = 3.51

iii. Range = 8.81 - 0.26 = 8.55, IQR = Q3 – Q1 = 3.51 – 0.94 = 2.57

iv. S^2= 4.69, S = 2.166

c. 1.5 (IQR) = 1.5 (2.57) = 3.855,
Q1 - 3.855 = -2.92 Lower limit
Q3 + 3.855 = 7.37 Upper limit

Two outliers whose values are 7.96 and 8.81. (Note that the whiskers end at 0.26 and 6.71)

```
                  -----------------
             |----I        +        I----------------------|     *     *
                  -----------------
           +---------+---------+---------+---------+---------+------minutes
         0.0  .26     1.6       3.2       4.8       6.4   6.7     8.0
```

2. a.

```
Leaf Unit = 1.0

    0 | 8
    1 | 0 5
    2 | 2 4 5
    3 | 0 5 5 8
    4 | 03
```

b. $X_{min} = 8$,
Q1 : (.25)(12) = 3rd value Q1 = (15+22)/2 = 18.5,
Q2 : (.5)(12) = 6th value Q2 = (25+30)/2 = 27.5,
Q3 : (.75)(12) = 9th value Q3 = (35+38)/2 = 36.5,
$X_{max} = 43$

c. $\bar{X} = 27.083$,   S = 11.735

d. i = (.4)(12) = .48
40th Percentile = the 5th ordered value = 24

3. a.

0	3
0	5 6 9 9
1	0 0 2 4 4
1	8
2	0 1 4
2	5
3	3

or

0	3 5 6 9 9
1	0 0 2 4 4 8
2	0 1 4 5
3	3

b. 16 (.3) = 4.8 30th percentile = 5th value = 9
16 (.7) = 11.2 70th percentile = 12th value = 20

c. i = .25(16) = 4th Q1 = (9+9)/2 = 9,
i= .75(16) = 12th Q3 = (20+21)/2 = 20. 5
IQR = 20.5 − 9 =11.5

1.5(IQR)=17.25

Q3 + 1.5(IQR) = 37.7   Upper limit

Yes, 42 days would be considered an outlier.

4. a.

i. $\bar{X}$ = \$3430,   Median = \$3500,   Mode = \$3500

ii. i = .25(20) = 5th value Q1 = (2200+2500)/2 = \$2,350,

i = .75(20) =15th value Q3 = (4300+4600)/2 = \$4,450

iii. Range = 5200 - 1000 = \$4,200;   IQR = 4450 - 2350 = \$2,100

iv. S^2 = 1,560,105;   S = \$1,249

v.

```
                         ----------------------------
         |----------------I             +           I-------|
                         ----------------------------
     +---------+---------+---------+---------+---------+-----
    800      1600      2400      3200      4000      4800
```

vi.

Compute: (1.5) IQR+Q3 = 7600.   No outliers

b. C.V. retainers = (1249/3430) *100 = 36.4%

C.V total compensation = (20,000/120,000) * 100 = 16.67%

No, each retainer is relatively more variable.

5. a.

	Fi	Cumulative	Cumulative relative
20-24	3	3	0.05
25-29	7	10	0.167
30-34	11	21	0.35
35-39	13	34	0.57
40-44	9	43	0.717
45-49	7	50	0.833
50-54	5	55	0.917
55-59	3	58	0.967
60-64	1	59	0.983
65-69	1	60	1

b. Right or positively skewed.

6. a. $\bar{X}$ = \$2491.25   S= \$282.11

$CV = (\frac{s}{\bar{x}} * 100)\% = 282.11/2491*100\% = 11.32\%$

b. The median = \$2405; 40th percentile: = \$2390

c. CV = 320/3600 = 0.0889 * 100 = 8.89%

No, they are less variable or smaller C.V.

7. a. $\bar{X}$ = 696.809 S = 483.188

b.

```
                 -------------
      |---------I       +    I------------|              *
                 -------------
      +---------+---------+---------+---------+---------+------
      0        400       800       1200      1600      2000
```

Q1: i = .25(21) = 5.25 = 6th value Q1 = 378

Q3: i = .75(21) = 15.75 = 16th value Q3 = 857

IQR = 857 - 378 = 479 IQR (1.5) = 718.5,

Lower limit = 378 – 718.5 = -340.5

Upper limit = 857 + 718.5 = 1575.5

Therefore, the data value 2063 is an outlier. (Note that the whisker ends at the data value 1434)

c. Point where 50% of the observations are more than \$684 and 50% are less

d. The proportion below $\bar{X}$ - 1S = 213.6 plus the proportion above $\bar{X}$ + 2S = 1663.2

You would expect: 0.34 + 0.475 = 0.815 or 81.5%

Actually proportion: out of 21, there are a total of 17 in the two regions giving 17/21 = 0.809 or 80.9%.

8. a.

Leaf Unit = 0.10

0	2455789
1	124589
2	3457889
3	13347
4	247
5	2568
6	238
7	689
8	
9	57

b.

$X_{minimum}$ = 0.2;

i = .25(40) = 10th

Q1 = (1.4+1.5)/2 = 1.45;

Q2 = (2.9+3.1)/2 = 3.0;

i =.75(40) = 30th

Q3 = (5.5+5.6)/2 = 5.55

IQR = 5.55 – 1.45 = 4.1

$X_{maximum}$ = 9.7

c.

```
                ---------------------
     |------I      +            I----------------------|
            ---------------------
     +---------+---------+---------+---------+---------+------
   0.0       2.0       4.0       6.0       8.0      10.0
```

IQR (1.5) = 6.15;
Lower limit = 1.45 – 6.15 = -4.7
Upper limit = 5.55 + 6.15 = 11.7
There are no outliers.

d. Distribution is positive or right skewed.

9. a. Cov (x,y) = -15.133 indicating a negative or an inverse linear relationship between price and demand

b. r = -15.133/ ((1.94) (8.214)) = - 0.9496, indicating a strong inverse or negative linear association between price and demand.

10. Nonsampling errors occur due to incorrect acquisitions of observations, improper selections samples, and the lack of response from individual in the sample. Examples are:

1. data acquisitions: wrong measurements on employees information (*answers may vary*)
2. exclusion of part-time employees from the sample (*answers may vary*)
3. missing information on employees characteristics; responses coming only from people who have strong feeling about the issues (*answers may vary*)

Probability

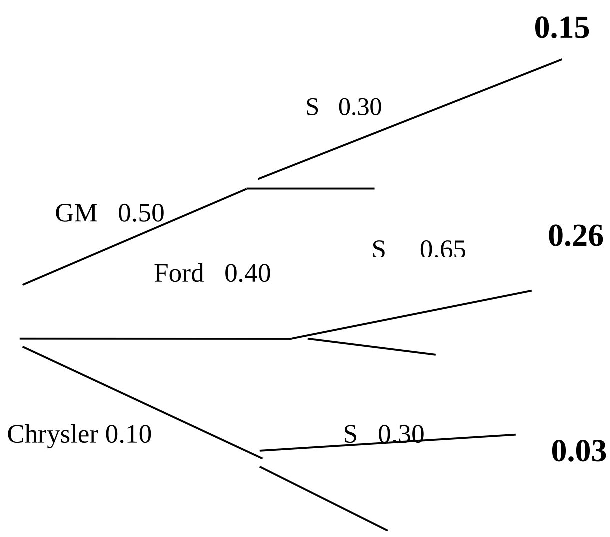
1.

	E	**U**	
M	.52	.08	**.60**
W	.28	.12	**.40**
	.80	**.20**	

Women Unemployed = .6 U
= .6 (.2) = .12
.3W = Women Unemployed
.3W = .12,
W = .4

a. P(W) = 0.40
b. P(U|M) = 0.08 / 0.60 = .133

2.



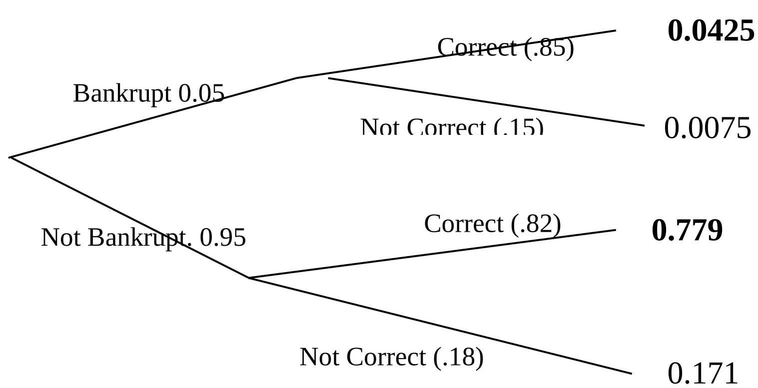
P (Strike) = .15 + .26 + .03 = .44

3.

a. (.9) (.9) = .81
b. (.1) (.1) = .01
c. (.9) (.1) + (.1) (.9) + (.9) (.9) = .99 or 1 – 0.01 = 0.99

4.

P (correct) = 0.0425 + 0.779
= 0.8215
0.8215 * 100 firms = 82 firms



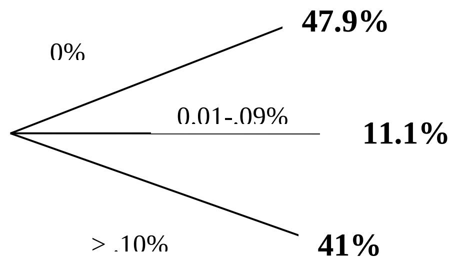
5.

a. P(<30 or 1^{st}) = 1 – (.07 + .09 + .05) = .79
b. P(repeat | < 30) = (.01+.02+.04)/(.09+.16+.11+.01+.02+.04)
= .16
c. P (≥ 30 | 1^{st}) = .36/.72 = .50
P (≥ 30 | repeat) = .21/.28 = .75; repeat offender is more likely to be ≥ 30

d. 50^{th} p of first offenders = .36 = 30
 a. 50^{th} p of repeat offenders = .14 = 35, repeat offenders are older

e. (.28) (.28) = .0784

6.



a. P (X≥ .01) = 11.1 + 41 = 52.1%

b. $P(X > 0.1 | X \geq 0.01)$ = .41/(.41+.111)
=.41 / .521 = .7869

7. a. i. (500-50-55-30)/500 = 365/500 = .73

ii. (60+50) = 110/340 = .3235

iii. 30/500 = .06

iv. $\frac{230+140-30}{500}$ = .68

b. Not mutually exclusive since P (work pt-time and < 6) = 60/500 = .12.
It will be zero if mutually exclusive.

Independent if P(work pt. time) * P(<6) = P(work pt. time and < 6)
Since 135/500 * 140/500 ≠ .12, the two are not independent.

8.

	P	S	A	
U	0.125	0	0.06	0.185
U^C	0.125	0.15	0.54	0.815
	0.25	0.15	0.60	1

a. P(U) = 0.185
b. P(A |U^c) = .54/.815 = .6625

9.

a.

Own Stock	City			
	A	B	C	Total
YES	85 (.17)	55 (.11)	50 (.10)	190 (.38)
NO	115 (.23)	95 (.19)	100 (.20)	310 (.62)
Total	200 (.40)	150 (.30)	150 (.30)	500 (1.0)

b. i. P(YES) = 0.38;

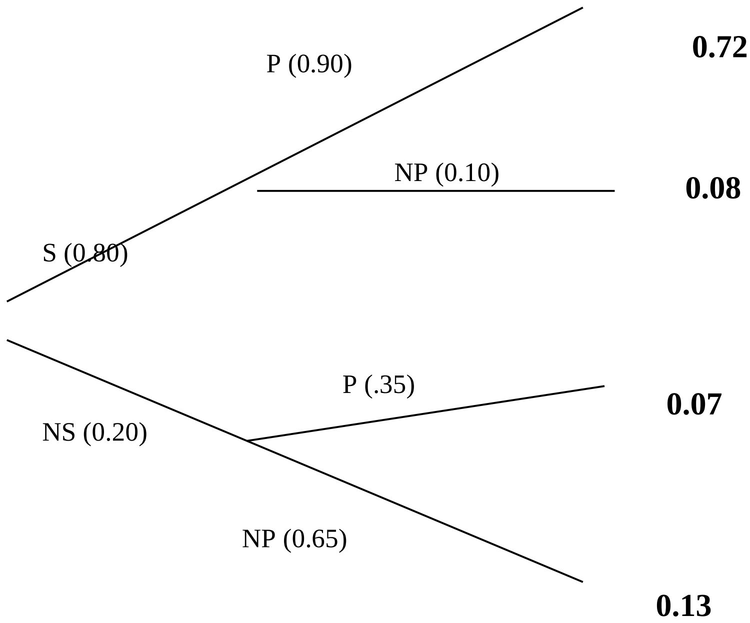
ii. P(YES and A) = 85/500 = 0.17;

iii. P(B and NO) = 95/500 = 0.19;

iv. P(B and YES) = 215/500 = 0.43

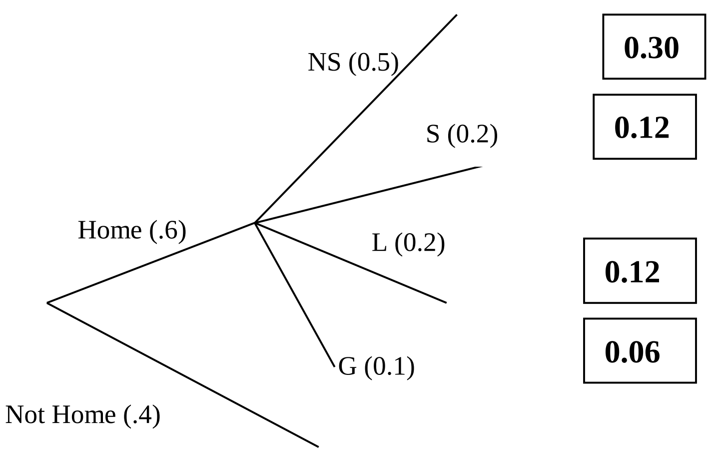
c. P (NO / NOT B) = (115+100) / (200+150) = 215/350 = 0.614

10.



a. P(S|P) = .72/ (.72+.07) = .72/.79 = 0.91
b. P(NS|NP) = .13/ (.08+.13) = 0.61 – –

11.



a. 0.4 + (.6) (.5) = .70
b. 0.12 + .06 = .18

12.

a.

Clerk	Defective (D)	Not Defective (ND)	
1	.008	.392	.40
2	.0075	.2925	.30
3	.0045	.2955	.30
	.02	.98	

b. P (1|D) = .008/.02 = .40

c. P (D or 2) = .02 + .03 - .0075 = .3125

d. P (D) = .02

e. P (2|D) = P (2) = .0075/.02 = .375 = .30; not independent

13. a.

P(M) = .54 => P(F) = .46;

P(U|M) = .074=> P(U and M) =.54*.074 = 0.04

P(U|F) = .0652=> P(U and F) = .46*.0652 =0.03

	E	U	
M	.5	.04	.54
F	.43	.03	.46
	.93	.07	1

b. P(F|U) = .03/.07 = .428

c. i. P(UM and UM) = (.04) (.04) = .0016

ii. P(1 or 2) = 1 – probability of none unemployed mail = 1 – (.96)(.96) = 0.784

Alternatively one can use binomial distribution and calculate

$P(X \geq 1|n=2, p = .04) = P(1)+P(2) = 2(.04)(.96) + (.04)^2 (.96)^0 = .0784$

d. P(U) = .07

e. P(Males) = .54   P(Male|U) = 0.04/0.07 = .571 $.54 \neq .571$

Therefore they are not independent.

(Can use any of the following formulas to check if the relation holds for the table)

*P(A and B) = P(A)*P(B)* $P(A|B) = \frac{P(AandB)}{P(B)}$

e.g. $P(U \text{ and } M) = 0.04 \neq P(U)*P(M) = (0.07)(.54) = .0378$

Discrete Probability Distribution / Binomial Distribution

1. a.

	1st	2nd	3rd	Total
Regular	.9 (.2)=0.18 $4.00	.7 (.3)=0.21 $2.00	.8(.5)=0.4 $0.50	0.79
Reduced	.1 (.2)=0.02 $2.00	.3 (.3)=0.09 $1.00	.2 (.5)=0.10 $0.25	0.21
Total	0.20	0.30	0.50	1

P (regular) = .79

b. $P(1^{st}|\text{regular}) = .18/.79 = .227$

c. μ = .18 (4) + .21 (2) + .4 (.50) + .02 (2) + .09 (1) + .10 (.25)
= $1.495

$$\mu_{regular} = \frac{.18(4)+.21(2)+.4(.5)}{.79} = \$1.696$$

d. $P(X \geq 2 \mid n = 8, p = .21) = 1 - \{P(0) + P(1)\}$

P(0) = (Combination of 8 and 0) $(.21)^0\ (.79)^8$ = .1517
P(1) = (Combination of 8 and 1) $(.21)^1\ (.79)^7$ = .3226

1 – (.1517 + .3226) = .5257

E (X) = np = 8 (.21) = 1.68; $\sigma = \sqrt{1.68(.79)} = 1.15$

e. $1.495 (75000) = $112,125 or $1.50 (75000) = $112,500

2. a. X = arrive on time p = .75
P(X = 4 | n = 4, p = .75) = .3164

b. P(X = 3) + P (X = 4) = .4219 + .3164 = 0.7383

3. .05 (2) + .15 (5) + .25 (10) + .35 (20) + .10 (50) + .10 (100)

$\mu = \sum Xi * P(X) = \$25.35,$ $\sigma^2 = \Sigma(X - \mu)^2 P(X) = 776.33$

4. a. $\mu = np = 12(.30) = 3.6\, people$

b. $P(X \geq 3 \mid n = 12, p = .30) = 1 - (P(0) + P(1) + P(2))$
$1 - (.0138 + .0712 + .1678) = .7472$

c. $P(X \geq 4 \mid n = 12, p = .30) = 1 - (P(0) + P(1) + P(2) + P(3))$
$1 - .4925 = 0.5075$

5. a. E(X) = .3 (100) + .2 (100) + .7 (100) = $120

b. Var (X) = $\sum (X - \mu)^2 P(x) = \$480.00$

6.

a. $\sum (X - \mu)^2 P(x) = \$480.00$

$P(X=3 \mid n = 15, p = .05)$ = (Combination of 15 and 3) $(.05)^3 (.95)^{12} = .0307$

b. $P(X \geq 3 \mid n = 15, p = .05) = 1 - (P(0) + P(1) + P(2)) = 0.0361$

c. If p is actually .05, then $P(X \geq 3)$ is very small (0.0361). Thus, it is unlikely to be actually 0.05.

7.

a. E(X) = μ = 10 (.2) = 2; standard deviation = $\sqrt{10(.2)(.8)} = 1.265$

b. $P(X \leq 2 \mid n = 10, p = .2) = .1074 + .2684 + .3020 = .6778$

c. $P(X=3 \mid n = 10, p = .8)$ = (Combination of 10 and 3) $(.8)^3 (.20)^7 = .000786$

d. p = P (pay cash| do not use store credit card) = .2/(1-.2) = 0.25
$P(X \leq 2 \mid n = 9, p = .25)$ =0 .60

8.

a. Mean = $\sum Xp(x) = 1.98$

b. Var(X) = $\sum (x - \mu)^2 p(x) = 1.919$ Standard deviation = 1.385

c. Expected Profit = 1.98 (2000) = $3,960

d. $P(X_1 + X_2 < 3) = P\{(0,0), (0,1), (0,2), (1,0), (1,1), (2,0)\}$
= (.15)(.15) + (.15) (.24) + (.15)(.31) + (.24) (.15) + (.24) (.24) + (.31) (.15) = .2451

9.

a. $P(X = 6 \mid n = 12, p = .30) = .0792$

b. $P(X \geq 6) = 1 - P(X \leq 5) = .1179$

c. $P(X \le 2) = P(0) + P(1) + P(2) = .2528$
d. $E(X) = np = 12 (.3) = 3.6$, or 3.6 (100) = $360

10.

a. $E(X) = 0 (.48) + 1 (.2) + 2 (.15) + 3 (.08) + 4 (.05) + 5 (.03) + 6 (.01) = 1.15$, $Var(X) = 2.1075$
b. $P(2) + P(3) = .15 + .08 = .23$
c. $P(X \ge 2) = 1 – (.48 + .20) = .32$
d. $P(X < 3) = P(0,0) + P(0,1) + P(0,2) + P(1,0) + P(1,1) + P(2,0) = .6064$

11.

a. $P(X \ge 3) = 1 – \{P(0) + P(1) + P(2)\} = 0 .74$
b. $P(2 \le X \le 5) = P(2) + P(3) + P(4) + P(5) = 0 .67$
c. $E(X) = 0 (.05) + 1 (.08) + 2 (.13) + 3 (.23) + 4 (.18) + 5 (.13) + 6(.08) + 7 (.06) + 8 (.06) = 3.78$
d. $\sigma = \sqrt{4.2} = 2.04$

12.

a. $E(X) = 0 (.1) + 1000 (.3) + 2000 (.4) + 3000 (.2) = \$1700$

b.

Cost	**Sales (x)**	**Net Profit**	**P(X)**	
3000*9.5 = 28500	0	-28500	.1	-2850
28500	1000 * 16	-12500	.3	-3750
28500	2000 * 16	3500	.4	1400
28500	3000 * 16	19500	.2	3900
	Expected Net Profit			**$-1300**

13.

a. $P(X \ge 9 \mid n = 15, p = .7) = 1 – P(X \le 8) = 1 - 0.1312 = 0.8688$
b. $P(X \ge 12 \mid n = 15, p = .7) = 1 – P(X \le 11) = 1 - 0.703 = 0.297$
c. $P(X \le 2 \mid n = 15, p = .25) = P(0) + P(1) + P(2)$
$= .0134 + .0668 + .1559 = 0.236$
d. $0.297 (250) = 74.25$

14.

a. P(1) + P(4) + P(5) = .25 + .15 + .10 =0 .5
b. 1 (.25) + 2(.33) + 3(.17) + 4(.15) + 5(.10) = 2.52
c. $\sigma = \sqrt{1.649} = 1.28$

15.

a. P(X=10 | n = 10, p = .6) = (Combination of 10 and 10) $(.6)^{10}$ $(.4)^{0}$ = .006
b. P(X≤2) = P(0) + P(1) + P(2) = .0001 + .0016 + .0106 = .0123

Normal Probability Distributions

1 a. $p(x \le 60) = p\left(z \le \frac{60\text{-}80}{10}\right) = p(z \le -2) = .0228$

b. $p(60 \le x \le 75) = p\left(\frac{60\text{-}80}{10} \le z \le \frac{75-80}{10}\right) = p(-2 \le z \le -0.5)$

c. $p(x \ge 90) = p\left(z \ge \frac{90\text{-}80}{10}\right) = p(z \ge 1) = 0.1587 \times 60 = 9.5 \approx 10 student$

2 a. $p(x \ge 4) = p\left(z \ge \frac{4-4.5}{0.6}\right) = p(z \ge -0.833) = 0.5 + 0.2967$

$$0.7967 \times 80 = 63.7 \approx 64$$

b. $10/80 = 0.125, \quad p(z \ge z_0) = .125 \to z_0 = 1.15$

$4.5 + 1.15(0.6) = 5.19$

c. $p\left(z \ge \frac{5.1\text{-}4.5}{0.6}\right) = p(z \ge 1)$ and $p\left(z \ge \frac{4.5\text{-}4.5}{0.6}\right)$

$p\left(\frac{z \ge 1}{z \ge 0}\right) = \frac{.1587}{.5} = 0.317$

3 a. $p(x>36.000)$ or $p(x<33.500) = p\left(z > \frac{36.000\text{-}34.265}{2000}\right) +$

$p\left(z < \frac{33.500 - 34.265}{2000}\right) = p(z > 0.861) + p(z < -0.383)$

$0.192 + 0.352 = 0.544$

b. $p\left(z > \frac{36.000\text{-}34.265}{2000}\right) \times p\left(z > \frac{36.000 - 35.185}{2400}\right)$

$p(z \ge 0.868) \times p(z > 0.34) = (.192)(.3669) = 0.07$

c. $p(x \le x_0) = .90 \quad 1.28 = \frac{x\text{-}35.185}{2400} \quad x=\$38.257$

4 a. $1\text{-}p(18 \le x \le 26) = 1\text{-}\left(\frac{18\text{-}20}{5} \le z \le \frac{26-20}{5}\right)$

$1 - p(-0.4 \le z \le 1.2) = 1 - (0.3849 + 0.1554) = 0.459$

b. $p(x \le 7) = p\left(z \le \frac{7\text{-}20}{5}\right) = -2.6 \quad p(z \le -2.6) = 0.0047$

$\frac{p(x \le 7)}{p(x \le 20)} = \frac{0.0047}{0.5} = 0.0094$

4. c. $p(x \geq 1) = 1 - p(x - 0)$

$$p(x = 0) = \frac{10!}{0!10!}(.54)^0(.46)^{10} = 0.000424$$

$1 - 0.000424 = 0.999576$

d. $p(x \leq 15) = p\left(z \leq \frac{15-20}{5}\right) =$ $P(z \leq -1) = 0.1587$

$= 0.1587 \times 200 = 31.74 \approx 32$

5. a. $p(x > 140.5) = p\left(z > \frac{140.5 - 140}{0.2}\right) = p(z > 2.5) =$

$.05 - 0.4938 = 0.0062$

b. $p(140.2 \leq x \leq 140.5)$ $z\left(\frac{140.2 - 140}{0.2}\right) = 1.0$

$p(1.0 \leq z \leq 2.5) = 0.4938 - 0.3413 = 0.1525$

c. $p(x < 139.8)$ *or* $p(x > 140.2)$

$p(z < -1)$ *or* p(z>+1)=0.5-0.3413=0.1587 × 2=0.3174

d. $p(x = 0 / n = 15,\ p = 0.3174) = \frac{15!}{0!15!}(0.3174)^0(0.6826)^{15} = 0.0033$

e. $p(x = 5 / n = 15,\ p = 0.3174) = \frac{15!}{5!10!}(0.3174)^5(0.6826)^{10} = 0.2110$

6. a. $p(x \leq 6) = z\left(\frac{6-12}{3}\right) = -2.0$ $p(z \leq -2) = 0.5 - 0.4772 = 0.0228$

b. $p(x > 15) = p\left(z > \frac{15-12}{3}\right) = p(z > 1.0) = 0.5 - 0.3413 = 0.158$

c. $z = \frac{x - \mu}{\sigma}$ $1.28 = \frac{x - 12}{3}$ x=15.84 ≈ 16weeks

7. a. p(x>214)=p$\left(z > \frac{214-250}{45}\right) = p(z > -0.8) = 0.2881 + 0.5 = 0.7881$

b. $p(x \leq 200) = p\left(z \leq \frac{200 - 250}{45}\right) = p(z \leq -1.11) = 0.5 - 0.3665 =$ 0.1335

Thus 13.35% of 52 weeks. That is, 0.1335 × $52 weeks = 6.9 \approx 7 weeks$

c. Find the 60th percentile, such that P(Z ≤ z*)= 0.6,

where z*= $\frac{x^*-250}{45}$. From table, z*=0.2533. Then, 0.2533= $\frac{x^*-250}{45}$,

and solving for x*=0.2533*45+250=261.4006