COMM 215: Business Statistics
Solution to Practice Problems 2

Sampling Distributions

1 a) $P(\bar{x} < 240) \text{ or } P(\bar{x} > 260) = P\left(z < \frac{240-250}{7.5/\sqrt{20}}\right) + P\left(z > \frac{260-250}{7.5/\sqrt{20}}\right)$

$P(z < -5.96) + P(z > 5.96) \approx 0$

b) $z = 1.04, \quad \frac{x-250}{7.5} = 1.04, \quad \Rightarrow \quad x = 257.8$

2 $\sigma/\sqrt{n} = 5.5/\sqrt{50} = 0.778 \qquad P(20 \le x \le 23) =$

$P\left(\frac{20-22}{0.778} \le z \le \frac{23-22}{0.778}\right) = P(-2.57 \le z \le 1.285) = 0.8964$

3 $P(\bar{p} \ge 0.10) = P\left(z \ge \frac{0.10-0.08}{\sqrt{\frac{(0.08)(0.92)}{400}}}\right) = 0.0708 \qquad \sigma_{\bar{p}} = 0.01356$

4 a) $\bar{p} = 15/100 = 0.15$

b) $P(\bar{p} \le .015) = P\left(z \le \frac{0.15-0.25}{\sqrt{\frac{(.25)(.75)}{100}}}\right)$

$P(z \le -2.31) = .5 - .4896 = .0104$

5 a) $P(\bar{x} \ge 26.2) = P\left(z \ge \frac{26.2-25}{3/\sqrt{36}}\right) = P(z \ge 2.4) = .5 - .4918 = .0082$

b) No, the probability is very small and is therefore very unlikely.

6 $P(\bar{p} \ge .75) = P\left(z \ge \frac{.75-.70}{\sqrt{\frac{(.70)(.30)}{200}}}\right) = P(z \ge 1.543) = 0.0618$

7 a) $\sigma_{\bar{p}} = \sqrt{\frac{(.10)(.90)}{400}} = 0.015$

b) $P(.09 \le \bar{p} \le .10) = P\left(\frac{.09-.10}{.015} \le z \le \frac{.10-.10}{.015}\right) =$

$P(-0.67 \le z \le 0) = 0.2486$

c) $P(\bar{p} < .08) = P\left(z < \frac{.08-.10}{.015}\right) = P(z < -1.33) =$

$.5 - .4082 = 0.0918$

Estimation and Hypothesis Testing

1 a) $H_0: \mu = 105$ $t_{.01,14d.f.} = -2.264$ $t = \dfrac{95.53 - 105}{15.39/\sqrt{15}} = -2.38$

$H_1: \mu < 105$

$-2.38 > -2.626$ do not reject H_0

insufficient evidence to conclude that mean < 105

b) $t_{.05,14d.f.} = -1.761$ $-2.38 < -1.761$ you would reject H_0

c) $.01<p<.025$

2 a) $n = \dfrac{(1.645)^2(.6)(.4)}{(.03)^2} = 721.6 = 722$

b) $\text{n}=\left(\dfrac{1.645(5)}{(.5)}\right)^2 = 270.6 = 271$

c) $110 \pm 1.96\left(6/\sqrt{70}\right)$, or 110 ± 1.406, or $(108.59, 111.406)$

3 a) $\begin{cases} H_0{:}\mu{=}10.8 \\ H_1{:}\mu{<}10.8 \end{cases}$ $t = \dfrac{10.2 - 10.8}{1.25/\sqrt{16}} = -1.92$; p-value: $0.025 < \text{p} < 0.05$

REJECT H_0 at α=.05 level of significance.

b) $\text{t}_{.025,15\text{d.f.}} = 2.131$ Since -1.92>-2.131 DO NOT REJECT H_0

4 type I : Approving a below standard shipment

type II: Refusing a shipment that is up to standard

5 a) $73 \pm 3.182\left(\dfrac{1.414}{\sqrt{4}}\right)$ 73 ± 2.25

b) $H_0: \mu = 77$

$H_1: \mu < 77$ reject H_0 if t<-2.353

$t = \dfrac{73 - 77}{1.414/\sqrt{4}} = -5.6577$

Reject H_0 and conclude that there is sufficient evidence to support the mean emission is less than 77 mg per cubic meter of exhaust.

6 a) H_0: $\mu = 5$ Reject H_0 if t >1.1812

H_1: $\mu > 5$

$$t = \frac{6.09 - 5}{6.41/\sqrt{11}} = 0.564. \quad \text{DO NOT Reject } H_0.$$

Insufficient evidence to support the airline's claim. p-value >.10

b) $6.09 \pm 2.228\left(\frac{6.41}{\sqrt{11}}\right)$ 6.09 ± 4.306 (1.784, 10.396)

7 $H_0 : p = 0.04$ Reject H_0 if z<-2.33

H_1 :p<0.04 $z = \frac{.10 - .04}{\sqrt{\frac{(.04)(.96)}{80}}} = 2.74$

Since 2.74 > -2.33, DO NOT reject H_0

No evidence that p< 0.04. p-value= p(z<2.74)=.9969

8 $H_0 : p = .51$

$H_1 : p \neq .51$

$p = \frac{690}{1335} = .5169$ Reject H_0 if $z > |1.645|$

$z = \frac{.5169 - .51}{\sqrt{\frac{(.51)(.49)}{1335}}} = .5044.$ DO NOT REJECT H_0. Insufficient evidence. Reject researcher's claim.

9 Assume standardized test scores follow a normal distribution with μ=100, σ=10.

$\begin{cases} H_0: \mu = 100 \\ H_1: \mu > 100 \end{cases}$ Test Statistic: $z = \frac{\bar{X} - 100}{10/\sqrt{n}}$; Rejection region: $Z > Z_{0.025} = 1.96$

with n=25, $\bar{X} = 103$, $z = \frac{103 - 100}{10/\sqrt{25}} = +1.5$. Since Z=1.5 <$Z_{0.025}$=1.96, do not reject at α=2.5%.

Insufficient evidence to support the claim of above average intelligence. p-value = 0.0668.

10 a) $n = \frac{(2.58)^2(.5)(.5)}{(.02)^2} \approx 4160.25$; n=4161

b) $0.48 \pm 2.58\sqrt{\frac{(.48)(.52)}{4161}}$, or 0.48 ± 0.01998, or (.46, .50)

11 a) $60 \pm 1.96\left(7.5/\sqrt{100}\right)$

$60 \pm 1.47 \rightarrow 58.53\%$; 61.47% true mean falls between 58.53% and 61.47% 95% of the time

b) $.07 \pm 1.96\sqrt{\frac{(.07)(.93)}{100}} \rightarrow .07 \pm .05 \rightarrow (0.02, 0.12) \rightarrow$ 95% confidence

True proportion of failing is between 2% and 12%

c) $n = \frac{(2.575)^2(7.5)^2}{(5)^2} = 14.92 = 15$ and $n = \frac{(2.575)^2(.07)(.93)}{(.1)^2} = 43.7 = 44$

To satify both conditions n must be at least 44.

12 a) *i*) $H_0: p = 0.25$ Reject H_0 if $z > |1.96|$

$H_1: p \neq 0.25$

$\hat{p} = \frac{15}{35} = 0.4286$ $z = \frac{.4286-.25}{\sqrt{\frac{(.25)(.75)}{35}}} = 2.44$

Since 2.44>1.96 reject H_0 and conclude that proportion using visual basic is different from 25%.

ii) p-value=$P(z \geq 2.44) \times 2 = .0073 \times 2 = 0.0146$

b) *i*) average cost > \$40,000 if average day > $\frac{40000-10000}{1200} = 25$

$H_0: \mu = 25$ Reject H_0 if $Z > Z_{0.05} = 1.645$

$H_1: \mu > 25$

$$z = \frac{27.2 - 25}{5.5/\sqrt{35}} = 2.37$$

Since Z=2.37 > $Z_{0.05}$=1.645 Reject H_0. Conclude average day >25.

The evidence is sufficient suggesting that mean will exceed \$40,000.

ii) p-value $P(z \geq 2.37) = .5 - .4911 = 0.0089$.

13 a) $\bar{x} = 237.829$ $s = 36.369$

$$237.83 \pm 1.645\left(\frac{36.37}{\sqrt{35}}\right) \qquad 237.83 \pm 10.11$$

(227.72, 247.94)

b) $\hat{p}$=more than 240 seconds $= 15/35 = 0.4285$

$$0.43 \pm 2.575\sqrt{\frac{(.43)(.57)}{35}}, \quad \text{or} \quad 0.43 \pm 0.2155, \text{ or } (0.2145, 0.6455)$$

14 $$n = \left(\frac{1.96(175.5)}{20}\right)^2 = 295.8 \rightarrow 296$$

15 $$n = \frac{(1.88)^2(.5)(.5)}{(.055)^2} = 292.099 \rightarrow 293$$

16 i) $\hat{p}=\frac{800-240}{800}=0.70 \quad \sigma_{\hat{p}}=\sqrt{\frac{(.70)(.30)}{800}}=0.0162$

$0.70\pm2.33(.0162)$ or 0.70 ± 0.0377 (.6623, .7377)

ii) $2.45\pm1.96\left(1.3/\sqrt{800}\right)$ $2.45\pm.09$ (2.36, 2.54)

17 $n=\left(\frac{2.575(300)}{45}\right)=294.65=295$

18 a) $18.30\pm t(.025, 8\text{ d.f})\left(6.3/\sqrt{9}\right)$

18.30 ± 4.8426 (\$13.46, \$23.14)

b) normally distrubuted population

c) $\hat{p}=\frac{49}{80}=0.6125$

$0.6125\pm2.17\sqrt{\frac{.6125(.3875)}{80}}$

$0.6125\pm.1182$

(0.4943, 0.7307)

d) $n=\left(\frac{(1.96)(6.30)}{1.00}\right)^2=152.47$, the sample size should be increased to 153.

19 a) $H_0:\mu=2.8$ Reject H_0 if $z<-1.645$

$H_1:\mu<2.8$ or if $t<-1.66$

$z=\frac{2.61-2.8}{0.9/\sqrt{100}}=-2.11$

Since $-2.11<-1.645$ Reject H_0 significant evindence that $\mu<2.8$ and therefore promotion was not profitable.

b) p-value $p(z<-2.11)=0.5-0.4826=0.0174$

or if t $.01<p<.025$

c) Type I since you are rejecting

d) $s/\sqrt{n}=.9/\sqrt{100}=.09$

reduced by half $=.09/2=.045$

$.9/\sqrt{n}=.045$ so $n=400$

Chi-Square Tests

1. $H_0 : p_1 = 0.50, \ p_2 = 0.25, \ p_3 = 0.15, \ p_4 = 0.10$

   H_1 : at least one p $\neq$ to specified values χ^2

f_i	e_i	χ^2
27	30	0.300
19	15	1.067
11	9	0.444
3	6	1.500

$\chi^2_{0.05,3} = 7.8147$. Since $3.311 < 7.8147$,

DO NOT REJECT H_0.

Brook's ideas regarding accounts are accurate. (No evidence to reject H_0)

2. H_0: Employment and wage rates are independent

   H_1: Employment and wage rates are NOT independent

	Yes	No	
High	18 (28.74)	40 (29.46)	58
Low	38 (27.26)	17 (27.74)	55
	56	57	113

Re ject if $\chi^2_{0.01,1} = 6.6349$

$\chi^2 = 4.016 + 3.945 + 4.235 + 4.16 = 16.355$

Since 16.355 > 6.6349, reject H_0 and conclude that classification are dependant.

3 a. H_0 : percentage in response categories are independent of type of firm.

   H_1 : percentage in response categories are NOT independent.

	Yes	Neutral	No	
US	50 (64.465	57 (46.290)	19 (15.237)	126
Foreign	60 (45.535)	22 (32.702)	7 (10.763)	89
	110	79	26	215

$\chi^2_{.10,2} = 4.605$. $\chi^2 = 16.06$

Since 16.06 > 4.605, Reject H_0 and conclude H_a. Evidence to indicate the two classification are dependant

b.

Sample proportion: $\bar{p} = {}^{50}\!/_{126} = .397$

The 90% confidence interval estimate for the percentage of U.S. firms that give hiring preferences to business majors with foreign language skills:

$$0.397 \pm 1.645\sqrt{\frac{(.397)(.603)}{126}}$$

$$0.397 \pm .072 \quad (.325, .469)$$

4.

$H_0: p_1 = p_2 = p_3 = p_4 = p_5 = 0.20$

H_1: at least one $p \neq 0.20$

e_i for each day $= 362 \times .20 = 72.4$, $\chi^2_{.05,4} = 9.48773$

Since $\chi^2 = 4.768 < \chi^2_{.05,4} = 9.48773$, DO NOT REJECT H_0.

Insufficient evidence to say absenteeism is higher on some days.

5.

H_0: preference is independent of experience (no relationship)

$H_{1:}$ preference is NOT independent of experience

$\chi^2_{.05,2} = 5.991$. Since $\chi^2 = 7.40136 > 5.991$, Reject H_0.

Sufficient evidence to conclude that preference and experiences are NOT independent. There is a relationship.

6.

a. H_0 : Scores are independent of gender

H_1 : Scores are NOT independent of gender

Reject H_0 if $\chi^2 > \chi^2_{.05,4} = 9.48773$ $\quad [(r-1)\times(c-1)=4]$

Test statistics: $\chi^2 = 1.172$ Since $1.172 < 9.487$

DO NOT Reject H_0 - Scores are independent of Genders

b. $$p(\bar{x} \leq 570) = p\left(z \leq \frac{570-550}{75/\sqrt{50}} \right) = 1.89$$

$p(z \leq 1.89) = 0.470610 + 0.5 = 0.9706$

.

7.

a. H_0: Same result obtained by ABC Inc.

H_1: Reason different than obtained by ABC Inc.

Reject H_0 if $\chi^2 > \chi^2_{.05,3} = 7.81473$

$\chi^2 = 1.0526 + 2.8125 + 2.3529 + 6.9231 = 13.1411$

Since $\chi^2 = 13.1411 > \chi^2_{.05,3} = 7.81473$,

reject H_0 and conclude that reasons are different.

b.

H_0: Reason for firing is independent to previous warning
H_0: Reason for firing is dependent to previous warning

Reject H_0 if $\chi^2 > \chi^2_{0.10,(4-1)(2-1)} = \chi^2_{0.10,3} = 6.251$.

Since $\chi^2 = \mathbf{5.423} < \chi^2_{0.10,3} = \mathbf{6.251}$, do not reject H_0.

8. a.

$$\begin{cases} H_0: \text{Finding a job is independent of working experience} \\ H_1: \text{Finding a job is NOT independent of working experience} \end{cases}$$

Reject H_0 if $\chi^2 > \chi^2_{.05,(r-1)(c-1)} = \chi^2_{.05,2} = 5.99$

Test statistics: $\chi^2 = 7.922$

Since $7.922 > 5.99$, Reject H_0.

Finding a job is NOT independent of work experience.

b.

$$\begin{cases} H_0: p_1=.70; p_2=.15; p_3=.15 \\ H_1: \text{at least one (or two) p is different} \end{cases}$$

Reject H_0 if $\chi^2 > \chi^2_{.10,2} = 4.60517$

Test statistics:

f_i	e_i	$\frac{(f_i - e_i)^2}{e_i}$
52	56	0.28571
16	12	1.33333
12	12	0.00000
80	80	1.61904

Since 1.61904 < 4.605, do not reject H_0; insufficient evidence that proportions differed from the special values.

Simple Linear Regression and Correlation

1 a) $\hat{y}=1.4235+0.53x$

b) $R^2=0821$. That is, 81.2% of the variation in ABC's rate of return is explained by the market of return.

c) $\begin{cases} H_0:\beta_1=0 \\ H_0:\beta_1\neq 0 \end{cases}$ $t=\frac{0.53}{0.103}=5.15$. Since 5.15 > 2.447, reject H_0 and conclude that model is significiant.

d) $\hat{y}=1.4235+0.53(5)=4.07$

$$4.07\pm 3.707(2.8225\sqrt{1+\frac{1}{8}+\frac{(5-2.5)^2}{752}} \Rightarrow 4.07\pm 11.137 \text{ or } (-7.067, 15.207)$$

2 a) $\hat{y}=10.548+0.00578x$

For each additional million dollars, the price per share increases by 0.00578 (in \$) while the initial price (constant) is \$10.55

b) $\begin{cases} H_0:\beta_1=0 \\ H_0:\beta_1\neq 0 \end{cases}$ $t=\frac{0.00578}{0.0039}=1.49$

p-value = $p(t\geq 1.49)\times 2 \Rightarrow (.05\times 2)\leq p\leq (.10\times 2)$, or $.10\leq p\leq .20$

So reject H_0 if $\alpha>.20$, but do not reject H_0 if $\alpha<.10$

c) $R^2=0.219$. That is, 21.9% of the variation in price per share is explained by the size of the offering.

d) $\hat{y}=10.548+0.00578(70)=10.95$. That is, the estimated mean price per share is \$10.95.

The 95% confidence interval estimate of the mean price per share for all companies with a size offering of \$70 million:

$$10.95\pm 2.306(.39)\sqrt{\frac{1}{10}+\frac{(70-64.3)^2}{10155.4}} \text{ or } 10.95\pm 0.29 \text{ or } (10.66, 11.24)$$

3 a) $\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$, where

$$SS_{xy} = 21115.07 - \frac{(92.93)(2725)}{12} = 12.2158$$

$$SS_{xx} = 720.22 - \frac{(92.93)^2}{12} = 0.554592$$

$$SS_{yy} = 619207 - \frac{(2725)^2}{12} = 404.917$$

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{12.2158}{0.554592} = 22.0267$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = \left(\frac{2725}{12}\right) - 22.0267\left(\frac{92.93}{12}\right) = 56.5048$$

$\hat{y} = 56.5048 + 22.0267x$, For each additional one percent increase in interest rates, the futures index increase by 22.0267 points.

b) $\begin{cases} H_0: \beta_1 = 0 \\ H_0: \beta_1 \neq 0 \end{cases}$ reject H_0 if $t > |t_{.005,\,10}| = 3.169$; test stat: $t = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}}$

$$S_{\hat{\beta}_1} = \frac{S}{\sqrt{SS_{xx}}}, \quad S = \sqrt{\frac{(SS_{yy} - \hat{\beta}_1 SS_{xy})}{(n-2)}} = \sqrt{\frac{404.917 - (22.0267)(12.2158)}{10}} = 3.68567$$

$$S_{\hat{\beta}_1} = \frac{3.68567}{\sqrt{0.554592}} = 4.94915, \qquad t = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}} = \frac{22.0267}{4.94915} = 4.45061$$

Since t=4.5061 > 3.169, reject H_0 at 5% level of significance and conclude that interest rate is a significant predict of futures index.

p value: .001 <p-value < .002

c) $$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}} = \frac{12.2158}{\sqrt{0.554592}\sqrt{404.917}} = 0.815180$$

The value of r suggests a strong positive correlation between interest rates and future index.

d) The 95% confidence interval estimate for the meanfutures index when interest rate is 7.8% :

$\hat{y}$ = 56.5048+22.0267(7.8) =228.313;

$$228.313 \pm 2.228(3.68567)\sqrt{\frac{1}{12}+\frac{(7.8-7.74417)^2}{0.554592}}$$

228.313±2.228(3.68567)(0.298252)

228.313±2.44915 or (225.864, 230.762)

Multiple Regression

1 a) Coefficient of x_2 = -0.021. The age at death will decrease by 0.021 years for every additional unit of cholestrol level.

b) $\begin{cases} H_0: \beta_1=\beta_2=\beta_3=0 \\ H_1: \text{at least one } \beta_i \neq 0 \end{cases}$ Test statistics: $F=\frac{MSR}{MSE}$

Rejection region: $F > F_{.05, 3, 36} = 2.84$

ANOVA table:

df	SS	MS	F
3	939	313	3.49
36	3227	89.64	
39	4166		

Since $F=3.49 > F_{.05, 3, 36} = 2.84$

Reject H_0 and conclude that model is significant in predicting lenght of life.

c) $\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$ Test statistics: $t = \dfrac{\hat{\beta}_1}{S_{\hat{\beta}_1}} \sim t_{0.005,36}$

Rejection region: $|t| > t_{0.005,36} = 2.724$

$$t = \frac{1.79}{0.44} = 4.068$$

Since t= 4.068 > 2.724, reject H_0 at 1% level of significance and conclude that average number of hours of exercise and age at death are linearly related.

d) $R^2 = 0.225$. That is, 22.5% of the variation in age at death is explained by x_1, x_2 and x_3.

e) $\hat{y}$ = 55.8 + 1.79 (8) - 0.021 (0.5) - 0.016 (10) = 69.95 years

f) x_1 = average hours of exercise per week because it has largest absolute value of t of 4.068.

2 a) $\hat{y}$=10+2.1(1.5)+13.6(3)=$53.95

b) $R^2 = \dfrac{90.400-43.912}{90.400} = 0.5142$. That is, 51.42% of the variation in store price is explained by variations in current dividents and rate of growth.

c) $\begin{cases} H_0: \beta_1 = \beta_2 = 0 \\ H_1: \text{at least one } \beta_i \neq 0 \end{cases}$ Test Statistic: $F = \dfrac{MSR}{MSE} \sim F_{0.05,2,7}$

Rejection region: $F > F_{0.05,2,7} = 4.47$

$$F = \frac{23244}{6273.143} = 3.7$$

Since $F = 3.7 < F_{0.05,2,7} = 4.47$, do not reject H_0. The model is not significiant.

d) 2.1(2.25) = $4.725 increase

3 a) $15232.5 + 2178.4x_1 + 7.8x_2 + 2675.2x_3 + 1157.8x_4$

b) for each additional room (x_1), the value of the house will increase by $ 2,178.4

c) DF for $\begin{cases} SSR & 4 \\ SSE & 25 \end{cases}$

d) $\begin{cases} H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1: \text{At least one } \beta_j \neq 0 \end{cases}$ Test statistic: $F = \dfrac{MSR}{MSE} \sim F_{0.05;4,25}$

Rejection region: $F > F_{0.05;4,25} = 2.76$

$$F = \frac{51060.72}{8235.60} = 6.2$$

Since $F = 6.2 > F_{0.05;4,25} = 2.76$, reject H_0 and conclude that the model is significant at 5% level.

e) $\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$ Test statistics: $t = \dfrac{\hat{\beta}_1}{S_{\hat{\beta}_1}} \sim t_{0.025,25}$

Rejection region: $|t| > t_{0.025,25} = 2.060$

$$t = \frac{2178.4}{778.0} = 2.8$$

Since $t = 2.8 > t_{0.025,25} = 2.060$, reject H_0 at 5% level and conclude that β_1 significant. The number of rooms (X_1) is an important predictor of value of a house (Y).

f) $R^2 = \dfrac{204,242.88}{410,132.88} = 0.49799.$

That is, 49.8% of the variation is the value of a house is explained by the 4 independent variables.

g) $\hat{y} = 15,232.5 + 2,178.4(9) + 7.8(7.500) + 2.675.2(2) + 1.157.8$
$= \$99,846.3$