Key Formula Sheet ASU ECN 221 ASWCC

Chapter 1: no key formulas.
Chapter 2: Relative Frequency=freq. of the class$/n$.
Approx. Class Width:
=(largest value-smallest value) /number of classes.
Chapter 3: sample and population means

$$\bar{x} = \sum x_i/n \text{ and } \mu = \sum x_i/N$$

Weighted mean and geometric mean

$$\bar{x} = \sum w_i x_i/w_i \text{ and } \bar{x}_g = [(x_1)(x_2)\ldots(x_n)]^{1/n}.$$

Interquartile Range: IQR $= Q_3 - Q_1$.
Population and sample variance

$$\sigma^2 = \frac{\sum(x_i-\mu)^2}{N} \text{ and } s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}.$$

Population and sample standard deviation

$$\sigma = \sqrt{\sigma^2} \text{ and } s = \sqrt{s^2}.$$

Coefficient of Variation

$$\left(\frac{\text{Standard deviation}}{\text{Mean}} \times 100\right)\%$$

z-Score: $z_i = \frac{x_i - \bar{x}}{s}$.
Population and Sample Covariance

$$\sigma_{xy} = \frac{\sum(x_i-\mu_x)(y_i-\mu_y)}{N} \text{ and } s_{xy} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n-1}$$

Population and Sample Pearson Correlation

$$\rho_{xy} = \sigma_{xy}/(\sigma_x\sigma_y) \text{ and } r_{xy} = s_{xy}/(s_x s_y).$$

Chapter 4: Counting Rule for Combinations

$$C_n^N = \binom{N}{n} = \frac{N!}{n!(N-n)!}.$$

Counting Rule for Permutations

$$P_n^N = n!\binom{N}{n} = \frac{N!}{(N-n)!}.$$

Probability Rules: $P(A) = 1 - P(A^C)$

Chapter 4 continued:

$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

$$P(A\cap B) = P(B)P(A|B) = P(A)P(B|A).$$

Multiplication Law for Independent Events

$$P(A\cap B) = P(B)P(A).$$

Bayes' Theorem

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)}$$

Chapter 5:
Discrete Uniform Probability Mass Function: $f(x) = 1/n$.
Expected Value of a Discrete R. V.: $E(x) = \mu = \sum x f(x)$.
Variance of a Discrete R. V.:

$$Var(x) = \sigma^2 = \sum(x-\mu)^2 f(x).$$

Number of Experimental Outcomes Providing Exactly x Successes in n Trials

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

Binomial Probability Mass Function

$$P(X=x) = f(x) = \binom{n}{x} p^x (1-p)^{(n-x)}.$$

Expected Value for Binomial Distribution: $E(x) = \mu = np$.
Variance for Binomial Distr.: $Var(x) = \sigma^2 = np(1-p)$.
Poisson Probability Mass Function:

$$P(X=x|\mu) = f(x) = \frac{\mu^x e^{-\mu}}{x!}.$$

Hypergeometric Probability Mass Function and Expected Value:

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \text{ and } E(x) = \mu = \frac{nr}{N}.$$

Chapter 5 continued: Variance for the Hypergeometric Distribution:

$$Var(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1-\frac{r}{N}\right)\left(\frac{N-n}{N-1}\right).$$

Chapter 6: Uniform PDF

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Normal PDF The density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Converting to the Standard Normal rv:

$$z = \frac{x-\mu}{\sigma}.$$

Exponential PDF and CDF for $x \ge 0$

$$f(x) = \mu^{-1}e^{-x/\mu} \text{ and } P(x \le x_0) = 1 - e^{-x_0/\mu}.$$

Chapter 7: expected value of $\bar{x}$

$$E(\bar{x}) = \mu.$$

Standard Deviation of $\bar{x}$ (Standard Error)

$$\sigma_x = \frac{\sigma}{\sqrt{n}}.$$

Expected Value and Std Dev (Standard Error) of $\bar{p}$

$$E(\bar{p}) = p \text{ and } \sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Finite Pop. Correction Factor: $\sqrt{(N-n)/(N-1)}$.
Chapter 8: Interval Estimate of Population Mean, σ known and unknown

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ and } \bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Necessary Sample Size for Interval Estimate of μ

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2}$$

Chapter 8 continued: Interval Estimate of p

$$\bar{p} \pm z_{\alpha/2}\frac{p(1-p)}{\sqrt{n}}$$

Necessary Sample Size for Interval Estimate of p

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2}$$

Chapter 9: Test Statistic for Hypothesis Tests About μ, σ known and unknown

$$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \text{ and } t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$$

Test Stat for Hypothesis About p

$$z = \frac{\bar{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Chapter 10: Point Estimate and Standard Error for Difference in Two Population Means

$$\bar{x}_1 - \bar{x}_2 \text{ and } \sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

Interval Estimate and Test Statistic for Difference in Two Means with Known Variances

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \text{ and } z = \frac{\bar{x}_1-\bar{x}_2-D_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

Interval Estimate and Test Statistic for Difference in Two Means with Unknown Variances

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \text{ and } t = \frac{\bar{x}_1-\bar{x}_2-D_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

Degrees of Freedom for t, Two Independent Random Samples

$$df = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Chapter 10 continued: Test Statistic (Matched Samples)

$$t = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$$

ANOVA Related:

$$\bar{x}_j = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j} \quad s_j^2 = \frac{\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2}{n_j-1} \quad \bar{\bar{x}} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} x_{ij}}{n_T}$$

$$MSTR = \frac{SSTR}{k-1} \quad \text{SSTR} = \sum_{j=1}^{k} n_j(\bar{x}_j-\bar{\bar{x}})^2 \quad MSE = \frac{SSE}{n_T-k}$$

$$SSE = \sum_{j=1}^{k}(n_j-1)s_j^2 \quad F = MSTR/MSE$$

$$\text{SST} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\bar{\bar{x}})^2 \quad \text{SST=SSTR+SSE}$$

Chapter 11: not covered in this course
Chapter 12: $y = \beta_0 + \beta_1 x + \epsilon$

$$E(y) = \beta_0 + \beta_1 x \quad \hat{y} = b_0 + b_1 x \quad b_0 = \bar{y} - b_1\bar{x}$$

$$b_1 = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} \quad SSE = \sum(y_i-\hat{y}_i)^2$$

$$SST = \sum(y_i-\bar{y})^2 \quad SSR = \sum(\hat{y}_i-\bar{y})^2 \quad \text{SST=SSR+SSE}$$

$$r^2 = \frac{SSR}{SST} \quad r_{xy} = (\text{sign of } b_1)\sqrt{r^2} \quad s^2 = MSE = \frac{SSE}{n-2}$$

Standard Error of the Estimate, $s = \sqrt{MSE}$.

$$\sigma_{b_1} = \frac{\sigma}{\sqrt{\sum(x_i-\bar{x})^2}} \quad s_{b_1} = \frac{s}{\sqrt{\sum(x_i-\bar{x})^2}} \quad t = \frac{b_1}{s_{b_1}}$$

For simple regression, $MSR = SSR$ because there is only one independent variable.

$$F = \frac{MSR}{MSE} \quad s_{\hat{y}*} = s\sqrt{\frac{1}{n}+\frac{(x^*-\bar{x})^2}{\sum(x_1-\bar{x})^2}}$$

Confidence Interval for $E(y^*)$: $\hat{y}^* \pm t_{\alpha/2}s_{\hat{y}*}$

$$s_{\text{pred}} = s\sqrt{1+\frac{1}{n}+\frac{(x^*-\bar{x})^2}{\sum(x_1-\bar{x})^2}}$$

Chapter 12 continued: Prediction Interval for y^*:

$$\hat{y}^* \pm t_{\alpha/2}s_{\text{pred}}$$

Residual for Observation i: $y_i - \hat{y}_i$
Chapter 13:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon$$

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p$$

$$SST = SSR + SSE \quad R^2 = \frac{SSR}{SST}$$

$$R_a^2 = 1 - (1-R^2)\frac{n-1}{n-p-1}$$

$$MSR = \frac{SSR}{p} \quad MSE = \frac{SSE}{n-p-1} \quad F = \frac{MSR}{MSE}$$

$$t = \frac{b_i}{s_{bi}}$$

Other Math Rule Reminders:

$$e^x = \exp(x)$$

$$\ln 1 = 0 \quad \ln e = 1$$

$$x! = (x)(x-1)(x-2)\cdots(2)(1)$$

$$0! = 1 \quad x^0 = 1$$

ECN221 Exam 1 A Fall 2015, ASU-COX

Choose the best answer. Do not write letters in the margin or communicate with other students in any way. If you have a question note it on your exam and ask for clarification when your exam is returned. In the meantime choose the best answer. Neither the proctors nor Dr. Cox will answer questions during the exam.

Please check each question and possible answers thoroughly as questions at the bottom of a page sometimes run onto the next page.

1. I have checked that my ID is bubbled in correctly. If it is bubbled in incorrectly I will get this question wrong. I also understand that questions and their possible answers may run onto the next page and so I should always check the top of the next page for possible answers. I understand that if I have a question I should simply make a note on my exam and ask Dr. Cox afterwards. I should always choose the best answer.

   (a) False.

   (b) I didn't read the directions and choose to get this question wrong.

   (c) True.

2. Suppose Terry has 10 potential dates for the weekend. However, Terry only wants 3 dates; one for each of Friday, Saturday and Sunday. Also, Terry refuses to go out with the same person more than once over the weekend. Given the pool of 10 potential dates how many possible combinations of dates are there for Terry?

   (a) 3628800.

   (b) 30.

   (c) 120.

   (d) A really big number not given here.

3. Imagine Dragons is giving 14 concerts in November 2015. The shows are all in Europe and their tour in Europe involves about 40 locations. How many different schedules are possible? That is, how many different possible combinations of cities are possible when giving 14 shows from a set of 40 cities? Choose the best answer.

   (a) over 1,000.

   (b) over 100,000.

   (c) over 1,000,000

   (d) over 1,000,000,000.

4. A stem and leaf display
   (a) is used to summarize categorical data.
   (b) is best used when there are a larger number of observations, e.g. $n \geq 1,000$.
   (c) is used to display either qualitative or quantitative data.
   (d) is used for quantitative data.

5. The relative frequency of "A"s in MAT211 is:

MAT211 Grade	frequency	relative frequency	percent frequency
A	166		
B	210		
C	124		
D	42		
E	22		

   (a) .294.
   (b) 29.4%.
   (c) 37.2%.
   (d) .22.

6. The cumulative percent frequency of "D"s in MAT211 is:

MAT211 Grade	cum freq	cum relative freq	cum percent freq
A	166		
B	376		
C	500		
D	542		
E	564		

   (a) 13.5%.
   (b) 7.4%.
   (c) 96.1%.
   (d) .961.

7. What type of variable is "Gender" in the data set below?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) quantitative
(b) categorical
(c) numeric
(d) ordinal

8. What is the mean income in the data set below?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	\$39,699,575	73	F
Kenny Chesney	\$32,956,240	45	M
Justin Timberlake	\$31,463,297	37	M
Bon Jovi	\$29,436,801	103	G
Rolling Stones	\$26,225,121	22	G

(a) \$32,006,788
(b) \$29,596,704
(c) \$30,954,207
(d) \$31,956,207

9. Income comes from concerts shown in the “Tour Dates” column and from other sources. Nevertheless, what was Taylor Swift’s average income per concert?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	\$39,699,575	73	F
Kenny Chesney	\$32,956,240	45	M
Justin Timberlake	\$31,463,297	37	M
Bon Jovi	\$29,436,801	103	G
Rolling Stones	\$26,225,121	22	G

(a) \$732,361
(b) \$543,830
(c) \$600,207
(d) \$274,485

10. Income comes from concerts shown in the “Tour Dates” column and from other sources. Nevertheless, which performer or group earned the most on a per concert basis?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) Taylor Swift

(b) Kenny Chesney

(c) Justin Timberlake

(d) Bon Jovi

(e) Rolling Stones

11. What is the median number of concerts (or "Tour Dates") in the data set?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) 37

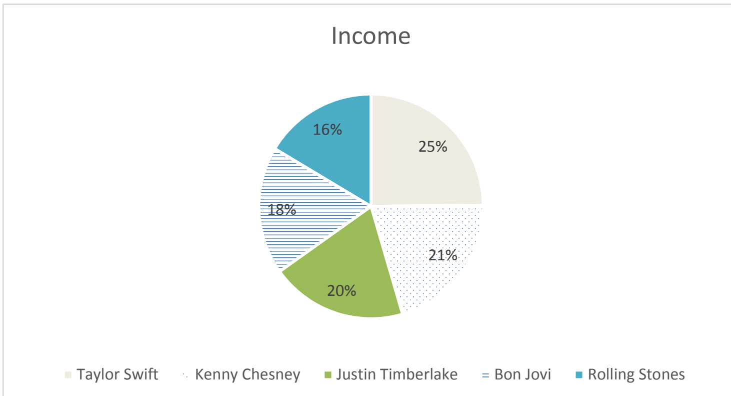
(b) 59

(c) 73

(d) 45

(e) 41

12. What type of graphical display is shown here?



(a) pie chart.

(b) bar chart.

(c) histogram.

(d) scatter plot.

13. What is the interquartile range for the income in the data set below?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) $4,972

(b) $2,492,455

(c) $3,519,439

(d) 50%

(e) 36

14. What is the standard deviation for the number of concerts (tour dates) in the data set below?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) 18.76
(b) 5.67
(c) 1034
(d) 32.16

15. What is the coefficient of variation for "Tour Dates"?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) 63.8%
(b) 174%
(c) 57.4%
(d) 32.2%

16. What is the range for the "Tour Dates" in the data set below?

Name	Income (2013)	Tour Dates	Gender
Taylor Swift	$39,699,575	73	F
Kenny Chesney	$32,956,240	45	M
Justin Timberlake	$31,463,297	37	M
Bon Jovi	$29,436,801	103	G
Rolling Stones	$26,225,121	22	G

(a) 81
(b) 103
(c) 56

(d) 29
(e) 36

17. A study by the Institute for Higher Education Policy found the values in the joint probability table below. The underlying data are for former college students that had taken out student loans. The table shows whether the student received a college degree versus whether they are successfully making their student loan payments. What is the probability that a former student with a loan is delinquent on their loan payments?

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

(a) .16.
(b) .34.
(c) .50.
(d) .42.

18. A study by the Institute for Higher Education Policy found the values in the joint probability table below. The underlying data are for former college students that had taken out student loans. The table shows whether the student received a college degree versus whether they are successfully making their student loan payments. What is the probability that a former student completed their degree?

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

(a) .58.
(b) .42.
(c) .26.
(d) .16.

19. A study by the Institute for Higher Education Policy found the values in the joint probability table below. The underlying data are for former college students that had taken out student loans. The table shows whether the student received a college

degree versus whether they are successfully making their student loan payments. What is the probability that a former student completed their degree and is currently satisfactory on their loan payments?

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

(a) .26.

(b) .42.

(c) .92.

(d) .68.

20. A study by the Institute for Higher Education Policy found the values in the joint probability table below. The underlying data are for former college students that had taken out student loans. The table shows whether the student received a college degree versus whether they are successfully making their student loan payments. What is the probability that a former student did not complete their degree given that they are currently delinquent on their loan payments?

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

(a) .34.

(b) .68.

(c) .50.

(d) .59.

21. I asked each of you how many cousins you have. Some answered with something like "a lot" or "too many to count." Of those that provided a number the average was 12.2 and the standard deviation was 38.5. The maximum was 1000. What is the z-value for the student with 1,000 counsins?

(a) 469.6.

(b) 25.6

(c) 5.0.

(d) 82.

22. What is the probability that two mutually exclusive events happen at the same time?

    (a) 1
    (b) .5
    (c) It depends on the original probabilities.
    (d) 0.
    (e) $P(A) + P(B) - P(A \cap B)$.

23. Chebyshev's inequality guarantees that less than 11.12% of the data will be more than 3 standard deviations away from the mean *no matter what distribution* as long as the mean and variance are finite. Using the interquartile range method for finding outliers, what is true? Hint: don't be afraid to draw pictures.

    (a) You could have more than 40% of your data be outliers.
    (b) The empirical rule still applies and you will see less than 1% of the data being outliers.
    (c) You can have at most 10% of the data be outliers; otherwise they are no longer "unusual."
    (d) We cannot say anything about the % of outliers we might observe unless we know something about the distribution.
    (e) You will have less than 11.12% of the data being outliers.

Key

1. c

2. c. Use the formula

$$\binom{10}{3} = \frac{10!}{3!7!} = 120.$$

You saw problems with combinatorics in homework 4 problems 3 and 4 among other places.

3. d. Use the formula

$$\binom{40}{14} = \frac{40!}{14!26!} = 23,206,929,840.$$

There is only one best answer in this case. You saw problems with combinatorics in homework 4 problems 3 and 4 among other places.

4. d. See the lecture notes for chapter 2, also see examples of stem and leaf displays in practice homework 2 #6-#8 and homework 2 #14 (wrong answer) and #18.

5. a. The relative frequency of "A"s in MAT211 is .294:

MAT211 Grade	frequency	relative frequency	percent frequency
A	166	.294	29.4
B	210	.372	37.2
C	124	.22	22
D	42	.074	7.4
E	22	.039	3.9

See the example in practice homework 2 #9 and other examples in lecture notes for chapter 2.

6. c. The cumulative percent frequency of "D"s in MAT211 is 96.1%:

MAT211 Grade	cum freq	cum relative freq	cum percent freq
A	166	0.294	29.4%
B	376	0.667	66.7%
C	500	0.887	88.7%
D	542	0.961	96.1%
E	564	1	100%

See the lecture notes for chapter 2 and practice homework 2 #10 and homework 2 #19.

7. b. categorical. This is straight from the definition. See the notes for chapter 1 and see homework 2 #4, 7, 10, 11, and 14 which all relate to understanding data and variable types.

8. d. $\bar{x}$ =\$31,956,207. Add all the incomes together and divide by $n = 5$. See chapter 2 lecture notes and the chapter itself. See also practice homework 1 #1, practice homework 3 #8 and homework 3 #10 and #24.

9. b. \$543,830 which is just \$39,699,575/73. See other examples of finding the average for help on this one.

10. e. Rolling Stones with \$1,192,051 per concert. Divide the total income \$26,225,121 by the total number of concert $n = 22$. Check lecture notes for chapter 3 on how to find the average.

11. d. 45. This is the middle value. See lecture notes for chapter 3 and practice homework 3 #4 and homework 3 #8.

12. a. this is a pie chart.

13. c. \$3,519,439 found from 32,956,240-29,435,801. You did problems with IQR for example on homework 3 #11. See also the lecture notes for chapter 3.

14. d. 32.16 after rounding. Use the formula in the notes and textbook. You calculated the standard deviation during class and for problem #9 on homework 3.

15. c.

$$\frac{32.156}{56} = .5742$$

We then convert this into a percentage to get 57.4%. This is a "B+" question. We did examples in class, see the lecture notes for chapter 3 when we calculated this for gasoline and car prices.

16. a. Take the max minus the min, $103 - 22 = 81$.

17. c. Take the value from the table.

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

18. b. Take the value from the table.

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

19. a. Take the value from the table.

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

20. b. Take the values from the table to make the calculations.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.34}{.5} = .68.$$

	holds a college degree		
loan status	yes	no	total
satisfactory	.26	.24	.50
delinquent	.16	.34	.50
total	.42	.58	1

21. b. Use the formula same as you used in homework 3 #15 and 16,

$$\frac{1000 - 12.2}{38.5} = 25.6.$$

22. d. 0, you simply need to refer back to rules in the notes and textbook.

23. a. This was a critical thinking or "A" question. You needed to remember the rule for finding outliers using the IQR method. Then you needed to do one of these:

    (a) Remember something like what happened when we added USC to our initial data set on tuition and recall that outliers don't impact the quartiles Q1 and Q3 so they don't impact the IQR and so everything above Q3 and below Q1 could be outliers using this method.

    (b) OR, draw a box and whisker plot such as I did in class and see that half the data could be outliers; hence my hint to not be afraid to draw a picture.

    (c) OR, you could have simply drawn a distribution or histogram with extremely heavy tails and "gaps" between the middle values and the tails and seen what that told you

    (d) OR, you could have had some other thought process or just been lucky in guessing.