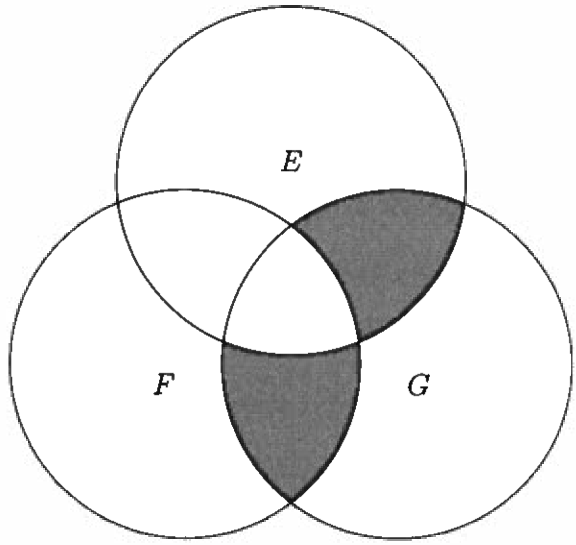
STATS 3Y03/3J04 Midterm #1
Greg Cousins, Sept 30, 2019

Name: ______________________________ **ID #:** ______________

- The test is 75 minutes long.
- The exam has questions on page 2 through 11; there are 18 multiple-choice questions printed on BOTH sides of the paper.
- Pages 12 to 20 contain no questions and can be used for rough work.
- You are responsible for ensuring that your copy of test is complete. Bring any discrepancies to the attention of the invigilator.
- Select the one correct answer for each question and enter that answer onto the answer sheet provided using an HB pencil.
- There are 18 multiple choice questions each worth 1 mark, and 1 question on correct computer card filling worth 1 mark (for a total of 19 marks)
- There is no penalty for a wrong answer.
- No marks will be given for the work in this booklet. Only the answers on the computer card (the scantron sheet) count for credit. You must submit this test booklet along with your answer sheet.
- You may use a Casio FX-991 MS or MS Plus calculator and there is a **formula sheet at the end of the test**; no other aids are permitted.
- **Good luck!!**

1. The shaded region in the diagram



represents which of the following?

(a) $(E \cup F) \cap G$ **(b)** $[G \cap (E \cup F)] \cap (F \cap E)'$ **(c)** $G \cap (E \cap F \cap G)'$
(d) $[E \cap (F \cup G)] \cap (G \cap F)'$

2. A garden is to be made with 2 rose bushes, 3 peonies, and 5 marigolds aligned in a row. How many unique garden arrangements are there?

(a) 5040 **(b)** 720 **(c)** 120 **(d)** 2520

Continued on page 3

3. A bin of 30 parts contains 3 that are defective. A sample of 4 parts is selected without replacement. What is the probability that that a sample contains at exactly two defective parts?

(a) 25% (b) 50% (c) 0.6% (d) 3.8%

4. A computer system uses passwords that contain exactly four characters and **which cannot have repeated characters**, with each character being either one of 26 lowercase letter (a-z), 26 uppercase letters (A-Z) or 10 integers (0-9). Assuming all passwords are equally likely, what is the probability that a password contains at least one integer?

(a) 0.456 (b) 0.723 (c) 0.217 (d) 0.515

Continued on page 4

5. A computer system uses passwords that contain exactly four characters **(repeated characters are allowed)**, and each character is either one of 26 lowercase letter (a-z), 26 uppercase letters (A-Z) or of 10 integers (0-9). Assuming all passwords are equally likely, what is the probability that a password contains one or two integers?

(a) 0.490 **(b)** 0.156 **(c)** 0.247 **(d)** 0.901

6. A batch of 500 deluxe blackboard erasers contains 15 that are defective. Three (3) are selected at random from the batch, without replacement. What is the probability that the third one selected is defective, given that the first one selected was defective, and the second one selected was okay?

(a) 6.89% **(b)** 1.37% **(c)** 5.11% **(d)** 2.81%

Continued on page 5

7. A lot of 100 semiconductor chips contains 20 that are defective. If two are selected at random, without replacement, what is the probability that the second one selected is defective?

(a) 0.3 (b) 0.2 (c) 0.4 (d) 0.5

8. Suppose that $P(A) = 0.47$, $P(B) = 0.32$, and that A and B are independent. What is $P(A \cup B)$?

(a) 0.150 (b) 0.790 (c) 0.639 (d) 0.239

Continued on page 6

9. A game of "guess the card" consists of one player selecting cards from a shuffled deck, and the other player, the "guesser", trying to randomly guess the card (e.g. "It's the ace of spades!"). A round consists of one cycle through the whole deck, and a player accumulates a point for each correct guess. Given that a standard deck has 52 cards, what is the probability that the guesser accumulates at least 2 points?

(a) 0.798 (b) 0.234 (c) 0.335 (d) 0.264

10. Suppose that X is a random variable with range $\{x_1, \ldots, x_n\}$ and cumulative distribution function $F(x)$. Which of the following statements are true:

I $F(x) = P(X \leq x)$,

II $\sum_{i=1}^{n} F(x_i) = 1$,

III $0 \leq F(x) \leq 1$ for all $x = x_1, \ldots, x_n$,

IV $F(x) = P(X = x)$.

(a) I and II (b) I, II, and III (c) I and III (d) II, III, IV

Continued on page 7

11. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. What is the probability that the third strike comes on the seventh well drilled?

(a) 0.038 (b) 0.455 (c) 0.054 (d) 0.049

12. Let X equal the number of typos on a printed page with a mean of $\lambda = 3$ typos per page. What is the probability that a randomly selected page has at least one typo on it?

(a) 0.857 (b) 0.950 (c) 0.750 (d) 0.542

Continued on page 8

13. Suppose we randomly select 5 cards without replacement from a well-shuffled deck of 52 cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

(a) 0.325 (b) 0.642 (c) 0.342 (d) 0.657

14. A survey shows that 73% of STATS 3Y03/3J04 students think that Greg is the nicest prof ever. If two students are picked at random, what is the probability that the first thinks that Greg is the nicest prof ever, and the second thinks there are probably nicer profs out there?

(a) 0.420 (b) 0.197 (c) 0.548 (d) 0.001

Continued on page 9

15. Suppose that 30% of McMaster undergraduate students are in their first year. President David Farrar likes to walk around and ask random students how they like the university. On average, how many students will Dr. Farrar need to talk to before he finds a first year?

(a)5.44 **(b)**6.00 **(c)**3.33 **(d)** 2.10

16. In 5 card poker, we say that a hand is *three of a kind* if the hand consists of three cards of a particular value, and two other cards. For example, $4\heartsuit, 4\diamondsuit, 4\clubsuit, K\diamondsuit, 7\spadesuit$ is three of a kind, with three 4's. Given a well-shuffled deck of 52 cards, what is the probability of being dealt three of a kind?

(a) 2.11% **(b)** 7.62% **(c)** 49.9% **(d)** 0.760%

Continued on page 10

17. A casino game consists of rolling two dice. If the outcome (the sum of the two outcomes) is a prime number P, you win \$$P$; if you roll anything else, you get nothing. What is the most you should pay to play one round of this game?

(a) \$2.00 **(b)** \$5.60 **(c)** \$2.56

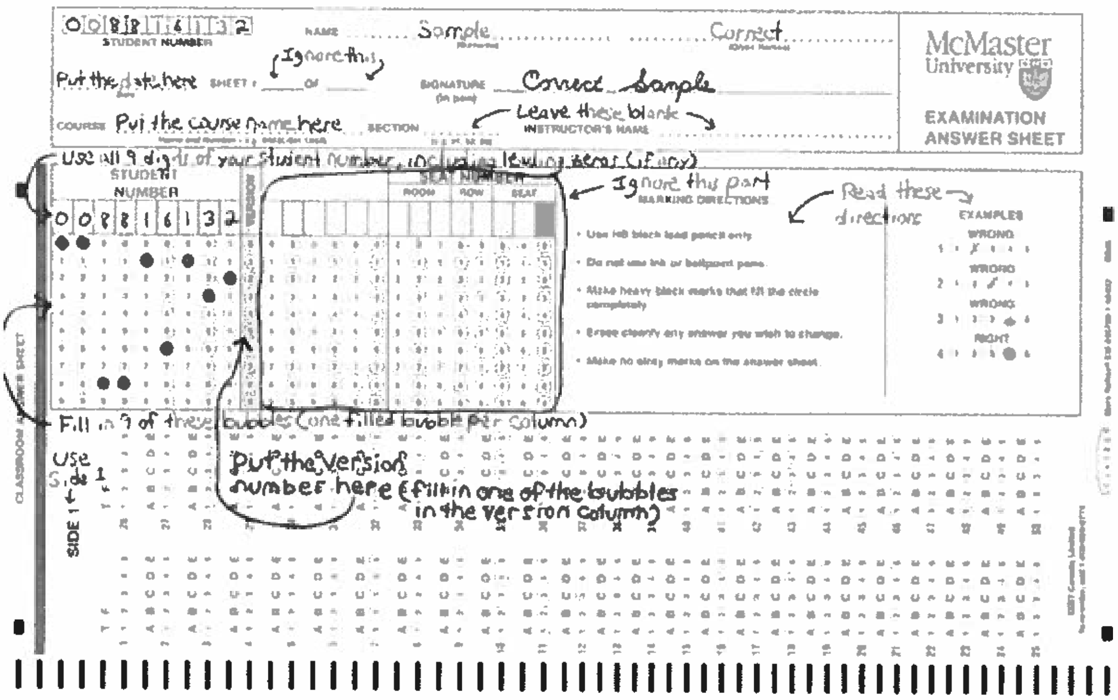
(d) \$0... Gambler's always lose in the long run!

18. Suppose we select 5 cards from an ordinary deck of playing cards. What is the probability of obtaining 2 or fewer hearts?

(a) 0.901 **(b)** 0.832 **(c)** 0.175 **(d)** 0.584

Continued on page 11

19. Correctly fill out the bubbles corresponding to all 9 digits of your student number, as well as the version number of your test in the correct places on the computer card. Note: You are writing **VERSION** 1. Hint:



END OF TEST QUESTIONS

Continued on page 12

CLASSROOM ANSWER SHEET

SIDE 1

1	F
2	D
3	D
4	D
5	A
6	D
7	B
8	C
9	D
10	C
11	D
12	B
13	A
14	B
15	C
16	A
17	C
18	A
19	
20	
21	
22	
23	
24	
25	
26–50	

STUDENT NUMBER

VERSION

SEAT NUMBER: ROOM ROW SEAT

McMaster University

EXAMINATION ANSWER SHEET

STUDENT NUMBER

NAME: Answer Key (Surname) (Given Names)

Date

SHEET # ___ OF ___

SIGNATURE: Answer Key

COURSE: Stats 3Y03/3J04 (Name and Number - e.g. ENGLISH 1A03)

SECTION (e.g. 01, 02, 03)

INSTRUCTOR'S NAME: Gregory Cousins.

MARKING DIRECTIONS

- Use HB black lead pencil only.
- Do not use ink or ballpoint pens.
- Make heavy black marks that fill the circle completely.
- Erase cleanly any answer you wish to change.
- Make no stray marks on the answer sheet.

EXAMPLES

1 WRONG

2 WRONG

3 WRONG

4 RIGHT

Extra page for rough work.

Continued on page 14

Extra page for rough work.

Continued on page 15

Extra page for rough work.

Continued on page 16

Extra page for rough work.

Continued on page 17

Extra page for rough work.

Continued on page 18

Formula Sheet (Page 1)

1. Addition Rule (events not mutually exclusive): $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

2. Addition Rule (events not mutually exclusive):

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

3. Addition Rule (mutually exclusive events): $P(A \cup B) = P(A) + P(B)$

4. Multiplication Rule (dependent events): $P(A \cap B) = P(A)P(B|A)$

5. Multiplication Rule (independent events): $P(A \cap B) = P(A)P(B)$

6. Conditional Probability: $P(B|A) = \frac{P(A \cap B)}{P(A)}$

7. Binomial: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$; $\binom{n}{x} = \frac{n!}{x!(n-x)!}$; $\mu = np$, $\sigma^2 = np(1-p)$

8. Negative Binomial: $f(x) = \binom{x-1}{r-1}(1-p)^{x-r}p^r$, $x = r, r+1, r+2, \ldots$; $\mu = \frac{r}{p}, \sigma^2 = \frac{r(1-p)}{p^2}$

9. Hypergeometric Distribution: $f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$, $\mu = np$, $\sigma^2 = np(1-p)\left(\frac{N-n}{N-1}\right)$

10. Poisson Distribution: $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0, 1, 2, \ldots$; $\mu = \sigma^2 = \lambda$

11. Normal Distribution: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$, $-\infty < x < \infty$

12. Exponential Distribution: $f(x) = \lambda e^{-\lambda x}$, $0 < x < \infty$; $\mu = \frac{1}{\lambda}$, $\sigma^2 = \frac{1}{\lambda^2}$

13. Marginal Distributions: $f_X(x) = \int f_{XY}(x, y)\, dy$, $f_Y(y) = \int f_{XY}(x, y)\, dx$

14. Correlation: $\rho_{XY} = \frac{\text{cov}(X,Y)}{\sqrt{V(X)V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

15. Transformation to standard normal: $z = \frac{X-\mu}{\sigma}$

16. Sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$

17. Outliers: $Q_1 - 1.5\,\text{IQR}$, $Q_3 + 1.5\,\text{IQR}$

18. Normal probability plot: $\Phi(z_j) = \frac{j-0.5}{n}$, $j = 1, 2, \ldots, n$

19. Central Limit Theorem formula: $z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}}$

20. z confidence interval for the mean: $\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Formula Sheet (Page 2)

21. t confidence interval for the mean: $\overline{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$

22. Confidence interval for a proportion: $\widehat{p} \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}$

23. z test for a mean: $Z_0 = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$; $\quad \beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$

24. t test for a mean: $T_0 = \frac{\overline{x}-\mu_0}{s/\sqrt{n}}$ **25.** z test for proportions: $Z_0 = \frac{X-np_0}{\sqrt{np_0(1-p_0)}} = \frac{\widehat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$;

$$\beta = \Phi\left(\frac{p_0-p+z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0-p-z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

Confidence interval for a difference in means:

26. Variances equal: $\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2,n_1+n_2-2}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$, $\quad s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

27. Variances unequal: $\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$, $\quad \nu = \left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2 \Big/ \left(\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}\right)$

28. t test for comparing two means (variances equal): $t_{n_1+n_2-2} = \frac{\overline{x}_1-\overline{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$

29. t test for comparing two means (variances unequal): $t_\nu = \frac{\overline{x}_1-\overline{x}_2}{\sqrt{(s_1^2/n_1)+(s_2^2/n_2)}}$,

30. Single variable Least Squares Regression line: $\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1\overline{x}$, $\widehat{\beta}_1 = \frac{s_{xy}}{s_{xx}}$, where

$$s_{xx} = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2 \text{ and } s_{xy} = \sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}$$

31. t-test for single variable regression: $t_{n-2} = \frac{\widehat{\beta}_1}{\sqrt{\widehat{\sigma}^2/s_{xx}}}$, where $\widehat{\sigma}^2 = \frac{SS_E}{n-2}$

32. Residual sum of squares: $SS_E = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$ **33.** Regression sum of squares: $SS_R = \widehat{\beta}_1 s_{xy}$

34. Total sum of squares: $SS_T = \sum_{i=1}^{n} y_i^2 - n\overline{y}^2$

35. Prediction Interval: $\widehat{y}_0 \pm t_{\alpha/2,n-2}\sqrt{\widehat{\sigma}^2\left(1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{s_{xx}}\right)}$

36. Sample correlation coefficient: $r = \frac{s_{xy}}{\sqrt{s_{xx}s_{yy}}}$ **37.** Coefficient of determination: $R^2 = \frac{SS_R}{SS_T}$

38. Total sum of squares: $SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{\cdot\cdot})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{N}$ (d.f. $= N-1$)

39. Error sum of squares: $SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{i\cdot})^2 = \sum_{i=1}^{a}(n_i-1)s_i^2$ (d.f. $= N-a$)

40. Treatment sum of squares: $SS_{\text{Treatments}} = \sum_{i=1}^{a} n_i(\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})^2 = \sum_{i=1}^{a}\frac{y_{i\cdot}^2}{n_i} - \frac{y_{\cdot\cdot}^2}{N}$ (d.f. $= a-1$)

41. Fisher's LSD Test: $\text{LSD} = t_{\alpha/2,N-a}\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$, **42.** Fisher's CI: $\overline{y}_{i\cdot} - \overline{y}_{j\cdot} \pm \text{LSD}$

END OF TEST PAPER