

Name: _____

Student Number: _____

E-mail Address: _____@mail.uoguelph.ca

**University of Guelph
Department of Mathematics and Statistics**

**STAT*2040: Statistics I
Test #1**

**VERSION 1
Red Computer Score Sheet**

Instructor: G. J. Umphrey

6 October, 2014
Time: 75 minutes

NOTES

1. This is a closed book examination, but one 8.5"×14" crib sheet (double sided) and a calculator are allowed.
2. This test consists of 20 equally-weighted multiple choice questions. Answer the multiple choice questions on the computer score sheet (it is also wise to circle your answers on the exam paper). There is no "guessing penalty" used in marking the multiple choice.
3. This test does not include any tables, as none are required.
4. Make sure you have a complete exam booklet!
5. Fill in the computer score sheet in pencil; make sure you include your name and student ID number, but you don't need a section number or email address. Fill in your name, student number and @mail.uoguelph.ca email user name at the top of this exam booklet.
6. **The entire exam booklet is to be handed in, including your rough work, along with the computer score sheet.**

1. Here's a basic "use your calculator efficiently" question: find the sample variance for the following data set

32.4, 42.7, 39.8, 37.0, 44.2, 43.3, 44.5, 41.6, 41.9

The answer is:

- (A) 3.70
 - (B) 3.92
 - (C) 13.66
 - (D) 15.37
 - (E) 40.82
2. An urn contains 10 balls: 2 red, 3 green, and 5 blue. Two balls are chosen randomly without replacement. What is the probability that both balls are the same colour?
- (A) $7/15$
 - (B) $14/45$
 - (C) $7/25$
 - (D) $7/30$
 - (E) None of the above.

Note: The next two questions are based on the following information.

Here is a data set consisting of blood fat concentration (in mg/dl) in the blood plasma of 25 men diagnosed with narrowing of the arteries.

184, 215, 221, 210, 208, 197, 250, 180, 212, 297, 168, 208, 180, 268, 219, 319, 250, 285, 221, 145, 168, 432, 92, 112, 87

The data was stored with the variable name “PlasmaContent” and the following analyses were performed with R:

```
> summary(PlasmaContent)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  87.0   180.0   210.0   213.1   250.0   432.0
> sd(PlasmaContent)
[1] 73.62479
> boxplot(PlasmaContent)
```

3. How many outliers will appear on the boxplot?
 - (A) 0
 - (B) 1
 - (C) 2
 - (D) 3
 - (E) 4 or more

4. Suppose 10 more observations were added to the data set. The mean of the 10 observations added is 217.6. What is the mean for the combined data set of 35 observations? (Pick closest value.)
 - (A) 213.9
 - (B) 214.4
 - (C) 214.9
 - (D) 215.4
 - (E) 215.9

5. Three fair dice are tossed. What is the probability that the numbers on the top faces of the three dice all differ?
- (A) $5/9$
 - (B) $5/12$
 - (C) $5/18$
 - (D) $5/36$
 - (E) None of the above.
6. A fair coin is flipped until the first head is obtained. If the first head occurs within the first four flips, what is the conditional probability that it occurred within the first two flips?
- (A) $2/3$
 - (B) $3/4$
 - (C) $4/5$
 - (D) $7/8$
 - (E) $15/16$

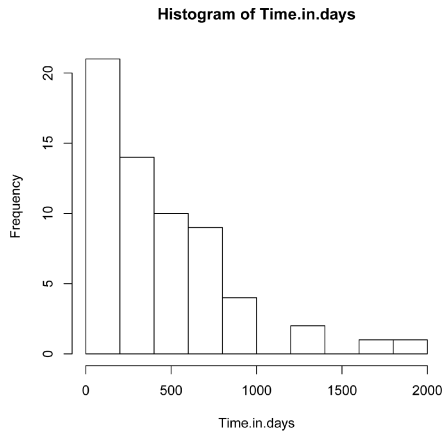
7. It is estimated that the percentages of A, B, O, and AB blood types in a certain population are, respectively, 40, 20, 30, and 10. Three individuals are picked (independently) from this population. (Assume that the population is very large). What is the probability that exactly one of the three individuals selected is Type A?
- (A) 0.072
 - (B) 0.144
 - (C) 0.216
 - (D) 0.432
 - (E) None of the above.
8. A tosses one fair coin and B tosses two fair coins. What is the probability that the number of heads obtained by A is equal to the number of heads obtained by B?
- (A) $\frac{1}{8}$
 - (B) $\frac{1}{4}$
 - (C) $\frac{3}{8}$
 - (D) $\frac{1}{2}$
 - (E) $\frac{5}{8}$

9. If $P(A') = 0.50$, $P(B) = 0.70$, and $P(A' \cap B) = 0.25$, then $P(B' | A') =$
- (A) 0.05
 - (B) 0.15
 - (C) 0.25
 - (D) 0.50
 - (E) None of the above.
10. A and B are two independent events. Let $P(A) = p$. If $P(B) = 4p$ and $P(A \cap B) = 0.160$, then the value of p is:
- (A) 0.04
 - (B) 0.08
 - (C) 0.20
 - (D) 0.40
 - (E) There is insufficient information to determine p .

11. Two events A and B have probabilities $P(A) = 3/5$ and $P(B) = 2/5$. Which one of the following statements is false?
- (A) If $P(A \cap B) = 0$, A and B are mutually exclusive events.
 - (B) If $P(A \cup B) = 1$, A and B are complements.
 - (C) If $P(B|A') = 0$, B is a subset of A.
 - (D) If $P(A \cap B) = 6/25$, A and B are independent events.
 - (E) If $P(A'|B) = 2/5$, A and B are not independent events.
12. A runner has six socks of a particular brand in her drawer. Two of the socks have holes. If she randomly grabs two of the six socks to wear, what is the probability that neither sock chosen has a hole?
- (A) $1/3$
 - (B) $2/5$
 - (C) $4/9$
 - (D) $8/15$
 - (E) None of the above.

NOTE: The next two questions refer to the following information.

A data set (in Hand et al.) consists of interval times (in days) between successive severe earthquakes for the years 1902–1977. There were 63 severe earthquakes, resulting in 62 interval times. A basic histogram plotted using R is shown.



13. Based on the histogram, it is most reasonable to conclude that:
- (A) the mean is greater than the median, because the distribution of interval times is skewed to the right.
 - (B) the median is greater than the mean, because the distribution of interval times is skewed to the right.
 - (C) the mean is greater than the median, because the distribution of interval times is skewed to the left.
 - (D) the median is greater than the mean, because the distribution of interval times is skewed to the left.
 - (E) the distribution of interval times will result in Chebyshev's Theorem being violated.
14. If one of the 62 interval times was randomly selected from the data set, the probability that the interval time selected is no greater than 400 days is closest to:
- (A) 0.23
 - (B) 0.34
 - (C) 0.56
 - (D) 0.73
 - (E) 0.87

15. The “Rare Disease” example discussed in class has an unusual outcome, at least in part because in the example the disease was rather rare. Now consider a similar problem for a more common disease that occurs in 10% of the population. For a randomly selected person the diagnostic test for the disease tests positive with probability 0.99 if the person has the disease but the probability of a positive test is only 0.02 if the person doesn’t have the disease. If the person tests positive, what is the probability they actually have the disease? (Pick closest value).
- (A) 0.55
 - (B) 0.65
 - (C) 0.75
 - (D) 0.85
 - (E) 0.95
16. If Bart’s grade of 74 on a test converted to a z-score of 1.2 and the mean score on the test was 56, the standard deviation on the test must have been:
- (A) 9
 - (B) 12
 - (C) 15
 - (D) 18
 - (E) None of the above.

17. Two fair dice (one red, one green) are tossed. Which pair of events are not mutually exclusive?
- (A) “red die shows a 4” and “red die shows an odd number”
 (B) “sum is odd” and “sum is even”
 (C) “green die shows a 1 or 2” and “sum is at least 9”
 (D) “both dice show the same number” and “red die shows a larger number than the green die”
 (E) “sum is at least 8” and “sum is no more than 8”
18. A biologist studying the behavioural ecology of lizards in a remote jungle in Papua New Guinea has captured, tagged and released 294 lizards, distributed by species and sex according to the following table.

	Species1	Species 2	Species 3	Total
Female	47	82	32	161
Male	35	84	14	133
Total	82	166	46	294

One of these 294 lizards is recaptured. Under the assumption that each lizard in the group has an equal probability of recapture, what is the probability the recaptured lizard belongs to Species 1 or 3, given its sex is female?

- (A) $79/294$
 (B) $79/161$
 (C) $79/128$
 (D) $128/294$
 (E) $161/294$

NOTE: The next two questions refer to the following information.

A frequency table was constructed for a sample of $n = 40$ observations. Two columns for the table are given below.

Class Limits	Frequency
12.5–17.5	4
17.5–22.5	8
22.5–27.5	11
27.5–32.5	10
32.5–37.5	7

19. The cumulative relative frequency for the class with midpoint 20 is:
- (A) 0.20
 - (B) 0.30
 - (C) 8
 - (D) 12
 - (E) None of the above.
20. The sample standard deviation for this grouped data is closest to:
- (A) 2.7
 - (B) 4.8
 - (C) 5.3
 - (D) 6.2
 - (E) 7.9