**Mgmt 230: Introductory Statistics**
**Instructor: Murray, James**
**Exam 1 Answers**
**Due Friday, February 20 at 11:00am. Late exams will earn zero credit.**

**Directions:** Answer the following questions on separate sheets of paper. All answers must be very neat and very clearly legible. If it is difficult to read an answer, no credit will be given. In order to earn the most partial credit possible, whenever using a formula, write down the formula first, then write down the formula with the numbers plugged in.

*You must complete this exam on your own without assistance from anyone.* Failure to comply with this condition is considered academic dishonesty and will result in the harshest penalty that Viterbo University will allow. You are allowed to use your notes, textbook, and computer.

1. Monthly data since 1996 for the average monthly percentage growth of two mutual funds is given in the Excel sheet `mutualfunds.xls`, posted on the class website. These numbers do not explicitly give information on dividends or capital gains payments, but this can give investors an idea of the growth rate and volatility of the earnings from these investments.

   (a) (3 points) What is the mean, median, and standard deviation of the return on the Janus Twenty Fund?
   $\bar{y} = 0.04$, $M_y = 0.93$, $s_y = 6.65$.

   (b) (3 points) What is the mean, median, and standard deviation of the return on the Schwab Core Equity Fund?
   $\bar{x} = -0.15$, $M_x = 0.18$, $s_x = 4.52$.

   (c) (5 points) What do your answers imply about the distribution of returns of each of these funds. Specifically, for each fund, is the distribution of returns skewed to the left, skewed to the right, or symmetric?
   `Both distributions are skewed to the left. This is what causes the mean return to be smaller than the median return.`

   (d) (5 points) What is the mean and standard deviation of the return on the portfolio: 50% on Janus Twenty and 50% on Schwab Core Equity?

$$
\begin{aligned}
&z = 0.5x + 0.5y \\
&\bar{z} = 0.5\bar{x} + 0.5\bar{y} = 0.5(0.04) + 0.5(-0.15) = -0.06 \\
&s_z^2 = 0.5^2 s_x^2 + 0.5^2 s_y^2 + 2(0.5)(0.5)s_{xy} \\
&= 0.25(44.17) + 0.25(20.46) + 2(0.5)(0.5)(25.19) = 28.752 \\
&s_z = \sqrt{s_z^2} = \sqrt{28.752} = 5.362
\end{aligned}
$$

`The mean of the portfolio is -0.06 and the standard deviation is 5.362.`

(e) (3 points) What is the covariance of returns on Janus Twenty and Schwab Core Equity? If using Excel, remember that Excel computes covariance by dividing by $n$ instead of $n-1$. Make sure to make the appropriate correction.
`= covar((B2:B153, C2:C153)*(152/151)) = 25.19.`

(f) (2 points) Does your answer above imply that the returns from these mutual funds move together or move away from each other.
`The returns of the mutual funds move together.`

(g) (3 points) A much better measure of how two variables are related is the correlation coefficient. The correlation coefficient is always a number between -1 and +1, where numbers close to zero imply little correlation, numbers close to +1 imply strong positive correlation, and numbers close to -1 imply strong negative correlation. The formula for the correlation coefficient is:

$$r_{xy} = \frac{s_{xy}}{s_x s_y},$$

where $r_{xy}$ is the correlation coefficient, $s_{xy}$ is the covariance, $s_x$ is the standard deviation of $x$, and $s_y$ is the standard deviation of $y$. Compute the correlation between Janus Twenty and Schwab Core Equity returns.

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{25.19}{(4.52)(6.65)} = 0.838$$

`The correlation is 0.838 which indicates strong positive correlation.`

2. Suppose for some population $\mu = 17$ and $\sigma^2 = 9$, and the population is normally distributed.

(a) (2 points) What is the z-score for $x_i = 20$? What is the probability of an observation having a value of 20 or below?

$$z = \frac{x_i - \mu}{\sigma} = \frac{20 - 17}{3} = 1.0$$

`P(z < 1.0) = normsdist(1.0) = 0.8413`

(b) (2 points) What is the z-score for $x_i = 25$? What is the probability of an observation having a value of 25 or above?

$$z = \frac{x_i - \mu}{\sigma} = \frac{25 - 17}{3} = 2.67$$

`P(z > 2.67) = 1 - normsdist(2.67) = 0.0038`

(c) (3 points) What is the probability of that a sample of size 15 yields a mean above 18?

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{18 - 17}{\frac{3}{\sqrt{15}}} = 1.29$$

`P(z > 1.29) = 1 - normsdist(1.29) = 0.0985`

(d) (3 points) What is the probability of that a sample of size 35 yields a mean above 18?

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{18 - 17}{\frac{3}{\sqrt{35}}} = 1.97$$

`P(z > 1.29) = 1 - normsdist(1.97) = 0.0244`

(e) (3 points) Your answers to the previous two questions are different because the sample sizes are different. Explain why?

`As the sample size increases, the variance of sampling distribution decreases. Therefore the sample mean is more likely to be close to the true population mean.`

3. Suppose there are two events, A and B, $P(A) = 0.5$, $P(B) = 0.6$, and $P(A \cup B) = 0.7$.

(a) (2 points) What is $P(A \cap B)$?

$$\begin{aligned} &P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ &0.7 = 0.5 + 0.6 - P(A \cap B) \\ &P(A \cap B) = 0.4 \end{aligned}$$

(b) (2 points) What is $P(A|B)$?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.4}{0.6} = 0.67$$

(c) (3 points) Are events A and B independent? Why or why not?

`They are not independent since` $P(A|B) \neq P(A)$.

(d) (3 points) Are events A and B mutually exclusive? Why or why not?
`No, because it is possible for both events to occur at the same time` ($P(A \cup B) = 0.7$).

(e) (3 points) What is $P(A \cap B')$?
`Use a Venn Diagram to show that:`
$P(A \cap B') = P(A) - P(A \cap B) = 0.5 - 0.4 = 0.1$.

(f) (2 points) What is $P(A \cap A')$?
`Impossible by definition.` $P(A \cap A') = 0$.

(g) (2 points) What is $P(A \cup A')$?
`Always happens.` $P(A \cup A') = 1.0$.

4. Suppose an Internet business makes customized business cards and letter head that it sells on its website. When an order is placed on the website, it must first produce the order, then ship the order. The average time it takes to produce the order is 2 days and the standard deviation is 0.5 days. The average time it takes to ship the order is 5 days with a standard deviation of 1.5 days. The time it takes to ship the order and the time it takes to produce the order are not correlated (zero covariance).

   (a) (2 points) What is the expected total amount of time from when the customer makes an order to the time she/he receives the order?
   `Z = X + Y`
   `E(Z) = E(X) + E(Y) = 2 + 5 = 7`

   (b) (3 points) What is the variance of the total time?

$$Var(Z) = Var(X) + Var(Y) + 2Cov(X,Y) = 0.5^2 + 1.5^2 + 2(0) = 2.5$$

5. (5 points) A statistical analysis of 1000 long distance phone calls found that the length of a long distance phone calls are normally distributed with a mean equal to 240 seconds and a variance equal to 1600 seconds. What is the probability that a randomly selected call lasts less than 180 seconds?

$$z = \frac{x_i - \bar{x}}{s} = \frac{180 - 240}{40} = -1.5$$

`P(z < -1.5) = 0.0668`

6. (5 points) What is the difference between Chebyshev's rule and the empirical rule? When should you use one over the other?

   `The empirical rule applies to data that is normally distributed. Chebyshev's rule is more general, and can be applied to any data, regardless of the distribution.`

7. (5 points) A restaurant survey asks for your opinions on a number of issues such as quality of service, quality of food, atmosphere of the restaurant, etc. For each question you have a choice of answering 1=very poor, 2=poor, 3=average, 4=good, 5=very good. The final question asks for your overall evaluation of the restaurant. What, if anything, is wrong with this survey?

   `The ordering of the questions influence the answer to the final question. For unbiased results, survey design and survey questions should not influence the outcomes of the survey.`

8. (5 points) A suspect on *CSI Miami* takes a lie detector test to try to prove his innocence in a murder investigation. In the lie detector test he says he did not kill anyone, but the lie detector test indicated this is a lie. The lie detector test has been proven to accurately identify lies 99.9% of the time. Comment on the validity of using this statistic to claim the suspect is guilty.

   `The accuracy probability is not the same conditional probability a court of law is interested in. Let A denote the event the lie detector test detected a lie. Let B denote the event the person is lying. The reported accuracy is P(A|B)=0.999. The probability the person is lying given the lie detector test signaled a lie is P(B|A), which could be very different.`

9. (5 points) Suppose two events are mutually exclusive. Can it be said whether these events are independent or dependent? Explain.

   The events must be dependent, since $P(A|B) = 0$ for mutually exclusive events, but in general $P(A) \neq O$. Since $P(A|B) \neq P(A)$, the events are dependent.

10. (Bonus 1 point) A bill from what historical currency decares that $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$?

    (a) French franc

    (b) **German mark!**

    (c) Israeli shekel

    (d) Marlboro lights

11. (Bonus 1 point) What famous Jewish rabbi and scholar is credited with an Isreali law that public schools cannot have larger classes than 40 students?

    (a) **Moses Maimondes!**

    (b) Jon Stuart Leibowitz

    (c) Yochanan ben Zakkai

    (d) Dalai Lama