STA 302f20 Assignment One[1]

Please do these review questions in preparation for Quiz One; they are not to be handed in. This material will not directly be on the final exam. Use the formula sheet on the course website.

1. The discrete random variable X has probability mass function $p(x) = |x|/20$ for $x = -4, \ldots, 4$ and zero otherwise. Let $Y = X^2 - 1$.
   (a) What is $E(X)$? The answer is a number. Show some work.
   (b) Calculate the variance of X. The answer is a number. My answer is 10.
   (c) What is $P(Y = 8)$? My answer is 0.30
   (d) What is $P(Y = -1)$? My answer is zero.
   (e) What is $P(Y = -4)$? My answer is zero.
   (f) What is the probability distribution of Y? Give the y values with their probabilities.

y	0	3	8	15
p(y)	0.1	0.2	0.3	0.4

   (g) What is $E(Y)$? The answer is a number. My answer is 9.
   (h) What is $Var(Y)$? The answer is a number. My answer is 30.

2. This question clarifies the meaning of $E(a)$ and $Var(a)$ when a is a constant.
   (a) Let X be a discrete random variable with $P(X = a) = 1$ (later we will call this a *degenerate* random variable). Using the definitions on the formula sheet, calculate $E(X)$ and $Var(X)$. This is the real meaning of the concept.
   (b) Let a be a real constant and X be a continuous random variable with density $f(x)$. Let $Y = g(X) = a$. Using the formula for $E(g(X))$ on the formula sheet, calculate $E(Y)$ and $Var(Y)$. This reminds us that the change of variables formula (which is a very big theorem) applies to the case of a constant function.

3. The discrete random variables X and Y have joint distribution

	$x = 1$	$x = 2$	$x = 3$
$y = 1$	3/12	1/12	3/12
$y = 2$	1/12	3/12	1/12

   (a) What is the marginal distribution of X? List the values with their probabilities.
   (b) What is the marginal distribution of Y? List the values with their probabilities.
   (c) Calculate $E(X)$. Show your work.
   (d) What is $Var(X)$? Show your work.
   (e) Calculate $E(Y)$. Show your work.
   (f) Calculate $Var(Y)$. Show your work. You may use Question 5a if you wish.

[1]Copyright information is at the end of the last page.

(g) Let $Z_1 = g_1(X,Y) = X + Y$. What is the probability distribution of Z_1? Show some work.

(h) Calculate $E(Z_1)$. Show your work.

(i) Do we have $E(X+Y) = E(X) + E(Y)$? Answer Yes or No. Note that the answer *does not require independence*, or even zero covariance.

(j) Let $Z_2 = g_2(X,Y) = XY$. What is the probability distribution of Z_2? List the values with their probabilities. Show some work.

(k) Calculate $E(Z_2)$. Show your work.

(l) Do we have $E(XY) = E(X)E(Y)$? Answer Yes or No.

(m) Using the well-known formula of Question 5b, what is $Cov(X,Y)$?

(n) Are X and Y independent? Answer Yes or No and show some work.

4. Let X_1 and X_2 be continuous random variables that are *independent*. Using the expression for $E(g(\mathbf{X}))$ on the formula sheet, show $E(X_1X_2) = E(X_1)E(X_2)$. Draw an arrow to the place in your answer where you use independence, and write "This is where I use independence." Because X_1 and X_2 are continuous, you will need to integrate. Does your proof still apply if X_1 and X_2 are discrete?

5. Using the definitions of variance covariance along with the linear property $E(\sum_{i=1}^{n} a_iY_i) = \sum_{i=1}^{n} a_iE(Y_i)$ (no integrals), show the following:

   (a) $Var(Y) = E(Y^2) - \mu_Y^2$

   (b) $Cov(X,Y) = E(XY) - E(X)E(Y)$

   (c) If X and Y are independent, $Cov(X,Y) = 0$. Of course you may use Problem 4.

6. Let X be a random variable and a be a constant. Show

   (a) $Var(aX) = a^2Var(X)$.

   (b) $Var(X + a) = Var(X)$.

7. Show $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$.

8. Let X and Y be random variables, and let a and b be constants. Show $Cov(X+a, Y+b) = Cov(X,Y)$.

9. Let X and Y be random variables, with $E(X) = \mu_x$, $E(Y) = \mu_y$, $Var(X) = \sigma_x^2$, $Var(Y) = \sigma_y^2$, $Cov(X,Y) = \sigma_{xy}$ and $Corr(X,Y) = \rho_{xy}$. Let a and b be non-zero constants.

   (a) Find $Cov(aX,Y)$.

   (b) Find $Corr(aX,Y)$. Do not forget that a could be negative.

10. Let $E(X_1) = \mu_1$, $E(X_2) = \mu_2$, $E(Y_1) = \mu_3$, $E(Y_2) = \mu_4$. Show $Cov(X_1 + X_2, Y_1 + Y_2) = Cov(X_1,Y_1) + Cov(X_1,Y_2) + Cov(X_2,Y_1) + Cov(X_2,Y_2)$.

11. Let $y_1, \ldots, y_n$ be numbers (not necessarily random variables), and $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$. Show

    (a) $\sum_{i=1}^n (y_i - \overline{y}) = 0$

    (b) $\sum_{i=1}^n (y_i - \overline{y})^2 = \sum_{i=1}^n y_i^2 \; - \; n\overline{y}^2$

    (c) The sum of squares $Q_m = \sum_{i=1}^n (y_i - m)^2$ is minimized when $m = \overline{y}$.

12. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be numbers, with $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$ and $\overline{y} = \frac{1}{n}\sum_{i=1}^n y_i$. Show $\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^n x_i y_i \; - \; n\overline{x}\,\overline{y}$.

13. Let $Y_1, \ldots, Y_n$ be independent random variables with $E(Y_i) = \mu$ and $Var(Y_i) = \sigma^2$ for $i = 1, \ldots, n$. For this question, please use definitions and familiar properties of expected value, not integrals or sums.

    (a) Find $E(\sum_{i=1}^n Y_i)$. Are you using independence?

    (b) Find $Var\left(\sum_{i=1}^n Y_i\right)$. What earlier questions are you using in connection with independence?

    (c) Using your answer to the last question, find $Var(\overline{Y})$.

    (d) A statistic T is an *unbiased estimator* of a parameter θ if $E(T) = \theta$. Show that $\overline{Y}$ is an unbiased estimator of μ.

    (e) Let $a_1, \ldots, a_n$ be constants and define the linear combination L by $L = \sum_{i=1}^n a_i Y_i$. What condition on the a_i values makes L an unbiased estimator of μ? Show your work.

    (f) Is $\overline{Y}$ a special case of L? If so, what are the a_i values?

    (g) What is $Var(L)$?

14. Here is a simple linear regression model. Independently for $i = 1, \ldots, n$, let $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where β_0 and β_1 are constants (typically unknown), x_i is a known, observable constant, and ϵ_i is a random variable with expected value zero and variance σ^2.

    (a) What is $E(Y_i)$?

    (b) What is $Var(Y_i)$?

    (c) Suppose that the distribution of ϵ_i is normal, so that it has density $f(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2\sigma^2}}$. Find the distribution of Y_i. Show your work. Hint: differentiate the cumulative distribution function of Y_i.

    (d) Let $\widehat{\beta}_1 = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}$. Is $\widehat{\beta}_1$ an unbiased estimator of β_1? Answer Yes or No and show your work.

15. Let $\mathbf{A} = \left(\begin{array}{rr} 2 & 5 \\ 1 & -4 \\ 0 & 3 \end{array} \right)$ and $\mathbf{B} = \left(\begin{array}{rr} 1 & 0 \\ 2 & 3 \\ -1 & 3 \end{array} \right)$ be matrices of constants. Which of the following are possible to compute? Don't do the calculations. Just answer each one Yes or No.

(a) $\mathbf{A}^{-1}$	(b) $\lvert\mathbf{B}\rvert$	(c) $\mathbf{A}+\mathbf{B}$
(d) $\mathbf{A}-\mathbf{B}$	(e) $\mathbf{AB}$	(f) $\mathbf{BA}$
(g) $\mathbf{A}'\mathbf{B}$	(h) $\mathbf{B}'\mathbf{A}$	(i) $\mathbf{A}/\mathbf{B}$

16. For the matrices of Question 15, calculate $\mathbf{A}'\mathbf{B}$. My answer is $\mathbf{A}'\mathbf{B} = \begin{pmatrix} 4 & 3 \\ -6 & -3 \end{pmatrix}$.

17. Let $\mathbf{c} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and $\mathbf{d} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. Verify that $\mathbf{c}'\mathbf{d} = 4$ and $\mathbf{c}\mathbf{d}' = \begin{pmatrix} 2 & 4 & -2 \\ 1 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$.

18. Which statement is true? Quantities in boldface are matrices of constants. Assume the matrices are of the right size.

    (a) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
    (b) $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{BA}+\mathbf{CA}$
    (c) Both a and b
    (d) Neither a nor b

19. Which statement is true?

    (a) $a(\mathbf{B}+\mathbf{C}) = a\mathbf{B}+a\mathbf{C}$
    (b) $a(\mathbf{B}+\mathbf{C}) = \mathbf{B}a+\mathbf{C}a$
    (c) Both a and b
    (d) Neither a nor b

20. Which statement is true?

    (a) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{AB}+\mathbf{AC}$
    (b) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$
    (c) Both a and b
    (d) Neither a nor b

21. Which statement is true?

    (a) $(\mathbf{AB})' = \mathbf{A}'\mathbf{B}'$
    (b) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$
    (c) Both a and b
    (d) Neither a nor b

22. Which statement is true?

    (a) $\mathbf{A}'' = \mathbf{A}$
    (b) $\mathbf{A}''' = \mathbf{A}'$
    (c) Both a and b
    (d) Neither a nor b

23. Suppose that the square matrices $\mathbf{A}$ and $\mathbf{B}$ are of the right sizes, and both have inverses. Which statement is true?
   (a) $(\mathbf{AB})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$
   (b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
   (c) Both a and b
   (d) Neither a nor b
24. Which statement is true?
   (a) $(\mathbf{A}+\mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
   (b) $(\mathbf{A}+\mathbf{B})' = \mathbf{B}' + \mathbf{A}'$
   (c) $(\mathbf{A}+\mathbf{B})' = (\mathbf{B}+\mathbf{A})'$
   (d) All of the above
   (e) None of the above
25. Which statement is true?
   (a) $(a+b)\mathbf{C} = a\mathbf{C} + b\mathbf{C}$
   (b) $(a+b)\mathbf{C} = \mathbf{C}a + \mathbf{C}b$
   (c) $(a+b)\mathbf{C} = \mathbf{C}(a+b)$
   (d) All of the above
   (e) None of the above
26. Let $\mathbf{A}$ be a square matrix with the determinant of $\mathbf{A}$ (denoted $|\mathbf{A}|$) equal to zero. What does this tell you about $\mathbf{A}^{-1}$? No proof is required here.
27. Recall that $\mathbf{A}$ symmetric means $\mathbf{A} = \mathbf{A}'$. Let $\mathbf{X}$ be an n by p matrix. Prove that $\mathbf{X}'\mathbf{X}$ is symmetric.
28. Matrix multiplication does not commute. That is, if $\mathbf{A}$ and $\mathbf{B}$ are matrices, in general it is *not* true that $\mathbf{AB} = \mathbf{BA}$ unless both matrices are 1×1. Establish this important fact by making up a simple numerical example in which $\mathbf{A}$ and $\mathbf{B}$ are both 2×2 matrices. Carry out the multiplication, showing $\mathbf{AB} \neq \mathbf{BA}$. This is also the point of Question 18.
29. Let $\mathbf{X}$ be an n by p matrix with $n \neq p$. Why is it incorrect to say that $(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{X}^{-1}\mathbf{X}'^{-1}$?
30. Let $\mathbf{A} = \left(\begin{array}{cc} 1 & 2 \\ 2 & 4 \end{array} \right)$ $\mathbf{B} = \left(\begin{array}{cc} 0 & 2 \\ 2 & 1 \end{array} \right)$ $\mathbf{C} = \left(\begin{array}{cc} 2 & 0 \\ 1 & 2 \end{array} \right)$
   (a) Calculate $\mathbf{AB}$ and $\mathbf{AC}$
   (b) Do we have $\mathbf{AB} = \mathbf{AC}$? Answer Yes or No.
   (c) Prove $\mathbf{B} = \mathbf{C}$. Show your work.

This assignment was prepared by Jerry Brunner, Department of Statistical Sciences, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The LaTeX source code is available from the course website: `http://www.utstat.toronto.edu/~brunner/oldclass/302f20`