The questions are practice for the quiz next week, and are not to be handed in. I would like you to bring in all of the code you used to run this assignment, and the output it generates. The best way to do this is just to run all of your code, and copy the console output to a .txt file and print it out. Clearly, your code should not print things that are not asked for, like an entire data frame. Even though you might want to look at the data frame while writing your code, your final version should only print the answers to the questions. Your console output (with code) should fit on two pages, double-sided, with two columns per side, at most. If your output is longer than 8 single pages (about 400 lines) you could probably tidy things up a bit. Remember, a whole lot of nonsense is not the same thing as concise, tidy code, even if the conclusions are the same.

If you want to experiment with R Markdown files to make it look nicer, be my guest! I might show these a little later in the course.

Remember, the computer assignments in this course are not group projects. You are expected to do the work yourself. You may compare numerical answers but do not share R code!

---

*This assignment was prepared by Craig Burkett with contributions from Jerry Brunner, both of the Department of Statistics, 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.
We’re going to be expressing all of our results using matrices shortly, so let’s get comfortable using them.

1. If the $p \times 1$ random vector $X$ has variance-covariance matrix $\Sigma$ and $A$ is an $m \times p$ matrix of constants, prove that the variance-covariance matrix of $AX$ is $A \Sigma A'$. Start with the definition of a variance-covariance matrix:

$$V(Z) = E(Z - \mu_z)(Z - \mu_z)' .$$

2. If the $p \times 1$ random vector $X$ has mean $\mu$ and variance-covariance matrix $\Sigma$, show $\Sigma = E(XX') - \mu \mu'$.

3. Let the $p \times 1$ random vector $X$ have mean $\mu$ and variance-covariance matrix $\Sigma$, and let $c$ be a $p \times 1$ vector of constants. Find $V(X + c)$. Show your work.

4. Let $X$ be a $p \times 1$ random vector with mean $\mu_x$ and variance-covariance matrix $\Sigma_x$, and let $Y$ be a $q \times 1$ random vector with mean $\mu_y$ and variance-covariance matrix $\Sigma_y$. Recall that $C(X, Y)$ is the $p \times q$ matrix $C(X, Y) = E ((X - \mu_x)(Y - \mu_y))'$.

   (a) What is the $(i,j)$ element of $C(X, Y)$?
   
   (b) Simplify for the special case where $\text{Cov}(X_i, Y_j) = 0$ for all $i$ and $j$.
   
   (c) Let $c$ be a $p \times 1$ vector of constants and $d$ be a $q \times 1$ vector of constants. Find $C(X + c, Y + d)$. Show your work.
   
   (d) If $p = q$, find an expression for $V(X + Y)$ in terms of $\Sigma_x, \Sigma_y$ and $C(X, Y)$. Show your work.

5. The joint moment-generating function of a $p$-dimensional random vector $X$ is defined as $M_X(t) = E \left( e^{t'X} \right)$.

   (a) Let $Y = AX$, where $A$ is a matrix of constants. Find the moment-generating function of $Y$ in terms of $M_X(t)$.
   
   (b) Let $Y = X + c$, where $c$ is a $p \times 1$ vector of constants. Find the moment-generating function of $Y$ in terms of $M_X(t)$.
   
   (c) Let $Y = AX + c$, as above. Find the moment-generating function of $Y$ in terms of $M_X(t)$.

6. Let $Z_1, \ldots, Z_p \overset{i.i.d.}{\sim} N(0, 1)$, and

$$Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_p \end{pmatrix} .$$

   (a) What is the joint moment-generating function of $Z$? Show some work.
(b) Let $Y = \Sigma^{1/2}Z + \mu$, where $\Sigma$ is a $p \times p$ symmetric non-negative definite matrix and $\mu \in \mathbb{R}^p$.

i. What is $E(Y)$?

ii. What is the variance-covariance matrix of $Y$? Show some work.

iii. What is the moment-generating function of $Y$? Show your work.

7. We say the $p$-dimensional random vector $Y$ is multivariate normal with expected value $\mu$ and variance-covariance matrix $\Sigma$, and write $Y \sim N_p(\mu, \Sigma)$, when $Y$ has moment-generating function $M_Y(t) = e^{t^\top \mu + \frac{1}{2} t^\top \Sigma t}$.

(a) Let $Y \sim N_p(\mu, \Sigma)$ and $W = AY$, where $A$ is an $r \times p$ matrix of constants. What is the distribution of $W$? Show your work.

(b) Let $Y \sim N_p(\mu, \Sigma)$ and $W = Y + c$, where $A$ is an $p \times 1$ vector of constants. What is the distribution of $W$? Show your work.

8. Let $Y \sim N_2(\mu, \Sigma)$, with

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

Using moment-generating functions, show $Y_1$ and $Y_2$ are independent.

9. Let $X = (X_1, X_2, X_3)'$ be multivariate normal with

$$\mu = \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$ 

Let $Y_1 = X_1 + X_2$ and $Y_2 = X_2 + X_3$. Find the joint distribution of $Y_1$ and $Y_2$.

10. Let $X_1$ be Normal($\mu_1, \sigma_1^2$), and $X_2$ be Normal($\mu_2, \sigma_2^2$), independent of $X_1$. What is the joint distribution of $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$? What is required for $Y_1$ and $Y_2$ to be independent? Hint: Use matrices.

11. Let $Y = X\beta + \epsilon$, where $X$ is an $n \times p$ matrix of known constants, $\beta$ is a $p \times 1$ vector of unknown constants, and $\epsilon$ is multivariate normal with mean zero and covariance matrix $\sigma^2 I_n$, where $\sigma^2 > 0$ is a constant. In the following, it may be helpful to recall that $(A^{-1})' = (A')^{-1}$.

(a) What is the distribution of $Y$?

(b) The maximum likelihood estimate (MLE) of $\beta$ is $\hat{\beta} = (X'X)^{-1}X'Y$. What is the distribution of $\hat{\beta}$? Show the calculations.

(c) Let $\hat{Y} = X\hat{\beta}$. What is the distribution of $\hat{Y}$? Show the calculations.
(d) Let the vector of residuals \( e = (Y - \hat{Y}) \). What is the distribution of \( e \)? Show the calculations. Simplify both the expected value (which is zero) and the covariance matrix.

12. Show that if \( X \sim N_p(\mu, \Sigma) \), \( Y = (X - \mu)'\Sigma^{-1}(X - \mu) \) has a chi-square distribution with \( p \) degrees of freedom. \text{Hint: See question 6}

13. Let \( X_1, \ldots, X_n \) be a random sample from a \( N(\mu, \sigma^2) \) distribution. Show \( \text{Cov}(X, (X_j - \overline{X})) = 0 \) for \( j = 1, \ldots, n \). This is the key to showing \( \overline{X} \) and \( S^2 \) independent, a fact you may use without proof in the next problem.

14. Recall that the chi-squared distribution with \( \nu \) degrees of freedom is just Gamma with \( \alpha = \frac{\nu}{2} \) and \( \beta = 2 \). So if \( X \sim \chi^2(\nu) \), it has moment-generating function \( M_X(t) = (1 - 2t)^{-\nu/2} \).

(a) Let \( Y = X_1 + X_2 \), where \( X_1 \) and \( X_2 \) are independent, \( X_1 \sim \chi^2(\nu_1) \) and \( Y \sim \chi^2(\nu_1 + \nu_2) \), where \( \nu_1 \) and \( \nu_2 \) are both positive. Show \( X_2 \sim \chi^2(\nu_2) \).

(b) Let \( X_1, \ldots, X_n \) be random sample from a \( N(\mu, \sigma^2) \) distribution. Show \( \frac{(n - 1)S^2}{\sigma^2} \sim \chi^2(n - 1) \).

\text{Hint: } \sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \overline{X} + \overline{X} - \mu)^2 = \ldots

15. Consider a two-factor analysis of variance in which each factor has two levels. Use this regression model for the problem:

\[ Y_i = \beta_0 + \beta_1 d_{i,A} + \beta_2 d_{i,B} + \beta_3 d_{i,A} d_{i,B} + \epsilon_i, \]

where \( d_{i,A} \) and \( d_{i,B} \) are dummy variables.

(a) Make a two-by-two table showing the four treatment means in terms of \( \beta \) values.

Use \textit{effect coding} (the scheme with 0, 1, -1). In terms of the \( \beta \) values, state the null hypothesis you would use to test for

i. Main effect of factor A

ii. Main effect of factor B

iii. Interaction

(b) Make a two-by-two table showing the four treatment means in terms of \( \beta \) values.

\textit{Use reference group coding.} In terms of the \( \beta \) values, state the null hypothesis you would use to test for the same three questions above.

(c) Which dummy variable scheme do you like more?
Arsenic is a powerful poison, which is why it has been used on farms for many years to kill insects. Even in very small amounts, arsenic can cause cancer in humans, and recently it has been found that rice and foods made from rice tend to be very high in arsenic. Brown rice is worse, by the way.

In a controlled experiment, pots of rice were prepared by either washing the rice first or not, and by cooking the rice in either a low, a medium or a high amount of water. The response variable is amount of arsenic in the cooked rice.

(a) Use a regression model with *cell means coding*. That’s the model with no intercept, and one indicator dummy variable for each treatment combination. You don’t have to say how the dummy variables are defined. That will become clear in the next part. Just give the regression equation.

(b) Write the expected amounts of arsenic in the table below, in terms of the \( \beta \) parameters of your model.

<table>
<thead>
<tr>
<th>Amount of Water</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Washed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unwashed</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(c) If you wanted to test whether the effect of washing the rice depended on how much water you cook it in, what is the null hypothesis? Give your answer in terms of the \( \beta \) values in your model.

(d) If you wanted to test whether washing the rice before cooking has any effect if the rice is cooked in a High amount of water, what is the null hypothesis? Give your answer in terms of \( \beta \) values.

(e) Suppose you want to test whether the amount of water used to cook the rice makes any difference if the rice has been washed. What is the null hypothesis? Give your answer in terms of \( \beta \) values.

(f) Averaging across different amounts of water used to cook the rice, does pre-washing affect the amount of arsenic in the rice. What null hypothesis would you test to answer this question? Give your answer in terms of \( \beta \) values.

(g) If you wanted to test whether the effect of the amount of water used to cook the rice depends on whether you wash it first, what is the null hypothesis? Give your answer in terms of \( \beta \) values.
17. Let’s return to the 2x2 Factorial design example from lecture, and recall that we were trying to see if the effect of Factor A depended on the levels of Factor B. You are working with some hapless analyst who just tried to do two separate t-tests (for Factor A) across the two subgroups of Factor B, and you need to convince them that they really should have just tested the interaction.

(a) Create a data frame with 40 students, 20 each of Male/Female, with half of each sex in Stats/Other programs. You should do this outside of your MC loop, as it doesn’t need to be repeated.

(b) Set the seed to ‘2015’ and do 10,000 MC simulations of the following:
   
i. Assign each student a score as follows: 55 + 5(if Female) + 15(if Stats) + N(0, 10). 10 is the SD, not variance
   
   To be clear, there is no interaction here, as the effects of Sex and Program are additive.

   ii. Do a two-sample t-test, checking if the levels of sex differ within each program, and get the p-value of the test. You have to do two tests; one for each program.

   iii. While you’re here, do the proper ANOVA with an interaction, and store the p-value.

(c) What proportion of the time did one t-test reject (at 5%), while the other did not? This proportion should be 5% (the Type I error rate) as there is no interaction.

(d) What proportion of the time was the ANOVA interaction (falsely) significant?

Give the hapless analyst a smack on the head!