STAT 666 Daily Quizzes

Q2. For each of the following statements about matrices A and B, specify "T=true," "Tsq=true, if we add that A & B are each square," or "F=false." (Assume that the quantity on the left exists.)

a. $|A+B| = |A| + |B|$
b. $(A+B)' = A' + B'$
c. $(A+B)^{-1} = A^{-1} + B^{-1}$
d. $tr(A+B) = tr(A) + tr(B)$
e. $|AB| = |A|\,|B|$
f. $(AB)' = A'\,B'$
g. $(AB)^{-1} = A^{-1}\,B^{-1}$
h. $tr(AB) = tr(A)\,tr(B)$

b, d, e (if A and B are square)

DIDN'T USE. For each type of analysis below, describe (i) the dimension of response **Y** and predictor **X** (if it exists), and (ii) the type of data found in **Y** and **X** (e.g., continuous, discrete numerical, categorical).

- Canonical correlation analysis
- Discriminant analysis (description)
- Classification analysis
- (Standard) confirmatory factor analysis
- Cluster analysis

Q3.

Let **x** be a random vector with mean **μ** and variance **Σ.** Find the mean and variance of $\mathbf{z} = \mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$

ACTIVITY with **x**, **y**, and **1** vectors:

(i) Find cor(**x**,**y**) for your two vectors…that correlation should be 0 if you're precise.

(ii) Find the length of $\bar{\mathbf{x}}$…that length—divided by sqrt(3)—should be the average of your three x's.

Q4. *(this may or may not be helpful on HW1…I don't even know why I mention that)*

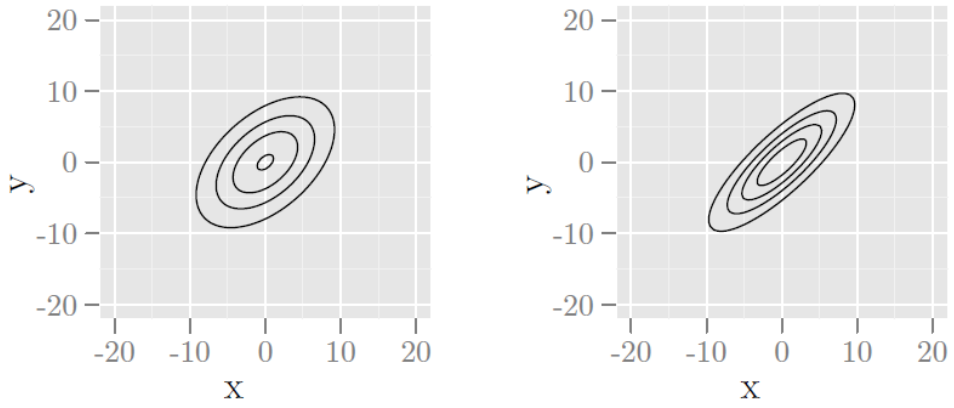
Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}$ be a random sample from a p-variate density. $\bar{\mathbf{x}}_n$ is the sample mean of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and $\bar{\mathbf{x}}_{n+1}$ is the sample mean of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}$. Show that:

$$\bar{\mathbf{x}}_n - \bar{\mathbf{x}}_{n+1} = \frac{1}{n+1}(\bar{\mathbf{x}}_n - \mathbf{x}_{n+1})$$

Q5.

The ACT math scores for Shadrach, Meshach, and Abednego are: 25, 26, and 21, respectively. The ACT reading scores for the three brothers had the same mean but were uncorrelated with the math scores. Shadrach scored a 25 on ACT reading. What are the ACT reading scores for Meshach and Abednego?

Q6.



(Fig 4.3 from RC)

$\sigma_{11} = \sigma_{22}$ for both plots

– Which has small $|\mathbf{\Sigma}|$ and which has large $|\mathbf{\Sigma}|$?

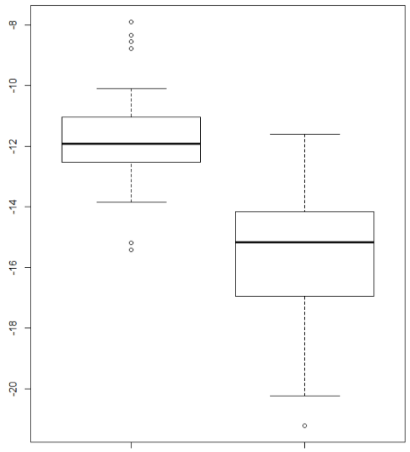
Q7. Give the fully specified distribution you would use to test the hypotheses below (assume normality for all data):

1. Compare the mean of (ACTmath, ACTread, ACTsci, ACTeng) for a sample of 100 students in this year's freshman class with the known mean from 2017.
2. After randomly assigning each of 20 pairs of twins to two different ACT prep classes, compare the mean ACT vector (ACTmath, ACTread, ACTsci, ACTeng) for the two training programs.
3. Every 10 minutes during the first week of Stat 641 last year, we measured the heart rate of each first year student for a total of 15 measurements per student. We are interested in comparing the mean heart rate profile (mean vector) for the 8 female students with the mean heart rate profile for the 7 male students.

Q8.

```
> xbar1    #noncarriers
     creatine.kinase            hemopexin lactate.dyhdrogenase      pyruvate.kinase
            43.05128             79.61538             12.53590            164.97436
> xbar2   # carriers
     creatine.kinase            hemopexin lactate.dyhdrogenase      pyruvate.kinase
           155.61765             94.00882             26.27059            247.50000
> Spl
                     creatine.kinase hemopexin lactate.dyhdrogenase pyruvate.kinase
creatine.kinase          12140.75953 -35.80727           1244.18093      2273.45847
hemopexin                  -35.80727 136.17631             -2.79539        46.80233
lactate.dyhdrogenase      1244.18093  -2.79539            210.76057       393.11177
pyruvate.kinase           2273.45847  46.80233            393.11177      2983.93626

> a <- solve(Spl) %*% (xbar1 - xbar2)
> a
                             [,1]
creatine.kinase      -0.00719236
hemopexin            -0.09924190
lactate.dyhdrogenase  0.01913890
pyruvate.kinase      -0.02314161
```



The boxplots to the right give plots of z=a'x for the two groups. Which boxplot is the one for carriers (left or right)?

Carriers on the right

Q9. Show that

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|}$$

is equal to

$$\prod_{i=1}^{s} \frac{1}{1+\lambda_i}$$

where λ_i, i=1,…,p, are the eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$

[Hint1: recall that if λ_i is an eigenvalue of **A**, then $1+\lambda_i$ is an eigenvalue of $\mathbf{I}+\mathbf{A}$

Hint2: if you're stumped, multiply by 1…or ask for another hint]

Q10.

Suppose that 40 customers from an online store are used in a repeated measures study. The 40 subjects comprise 10 each from 4 different customer types (I, II, III, and IV). During the study, each customer uses each of 3 website designs (A, B, and C) and the number of clicks used to navigate to a purchase was recorded. Consider a traditional univariate split plot/repeated measures design. Give the ANOVA table with Source, df, and F-statistic described for each term in the model.

Q11.

Some types of software (e.g. R's "eigen" function) require a symmetric matrix to obtain non-imaginary e'values and e'vectors. $\mathbf{E}^{-1}\mathbf{H}$ is not symmetric. We can accommodate "eigen" by simply calculating the spectral decomposition of $\mathbf{E}^{-1/2}\mathbf{H}\,\mathbf{E}^{-1/2}$. Show that the e'values of $\mathbf{E}^{-1/2}\mathbf{H}\,\mathbf{E}^{-1/2}$ are the same as the eigenvalues of $\mathbf{E}^{-1}\mathbf{H}$, but that the e'vectors of $\mathbf{E}^{-1/2}\mathbf{H}\,\mathbf{E}^{-1/2}$ are of the form $\mathbf{E}^{1/2}\,\mathbf{x}$, where $\mathbf{x}$ is an e'vector of $\mathbf{E}^{-1}\mathbf{H}$. (Of course, assume that E is nonsingular.)

[Hint: start with $(\mathbf{E}^{-1}\mathbf{H} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$]

NOPE. Suppose that 40 customers from an online store are used in a repeated measures study. The 40 subjects comprise 10 each from 4 different customer types (I, II, III, and IV). During the study, each customer uses each of 3 website designs (A, B, and C) and the number of clicks used to navigate to a purchase was recorded. Devise a univariate (over-simplified) approach for a permutation test of the following in a repeated measures analysis:

- Between subjects factor (e.g., customer type)
- Within subjects factor (e.g., website design)

FYI, the preamble of the exam:

Stat 666 - Fall 2019
Dr. William Christensen
Midterm Exam — October 17-23, 2019

NAME {2 pts.}: _____

NOTES: (1) Throughout the exam, use numbers in your answer wherever possible. For example, suppose the answer to a question is npk, where it can be deduced that $p = 3$ and $k = 5$, but where n is unknown. An answer of "$15n$" will receive full credit whereas answers of "npk" or "$5pn$" will receive partial credit. (2) There is a blank page attached that you may use for answers if necessary; leave a note such as "see #3 on attached page."

The following may be helpful at some point: $\begin{bmatrix} a & c \\ c & b \end{bmatrix}^{-1} = \frac{1}{ab-c^2}\begin{bmatrix} b & -c \\ -c & a \end{bmatrix}$.

Q12. A PhD program gathers information on students' grades (final %) in each of their classes and is interested in comparing across groups of their PhD students by their most recent degree: (i) MS in Stat (6 students), (ii) BS in Stat (30 students), (iii) BS in Math (44 students), (iv) other (20 students). On each student, the following are gathered:

Theory	Theory	Theory	Meth.	Meth.	Meth.	Comp.	Comp.	Comp.	Writ	Writ	Writ
Sem1	Sem2	Sem3	Sem1	Sem2	Sem3	Sem1	Sem2	Sem3	Sem1	Sem2	Sem3

1. Describe the test for the subject matter x semester interaction, including: the form of the test statistic, any contrast matrices needed, and the distribution of the test statistic.
2. What if you are specifically interested in how grades change across the 3 semesters when comparing Theory with Non-theory? Give : the form of the test statistic, any contrast matrices needed, and the distribution of the test statistic.
3. Does the Theory vs. Non-theory comparison change across the most-recent-degree groups? Give : the form of the test statistic, any contrast matrices needed, and the distribution of the test statistic.