STAT 3690 Homework 1

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Due at Feb 28 11:59 am (Central Time)

Answers must be submitted electronically via Crowdmark. Please enclose your R source code (if applicable) as well.

1. The function $\mathrm{cov}(\cdot,\cdot)$ is bilinear, i.e., for random vectors $\mathbf{W}$, $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ and fixed matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, one has $\mathrm{cov}(\mathbf{AW}+\mathbf{BX},\mathbf{Y}) = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{WY}} + \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{XY}}$ and $\mathrm{cov}(\mathbf{W},\mathbf{CY}+\mathbf{DZ}) = \boldsymbol{\Sigma}_{\mathbf{WY}}\mathbf{C}^\top + \boldsymbol{\Sigma}_{\mathbf{WZ}}\mathbf{D}^\top$, where $\mathbf{AW}+\mathbf{BX}$ and $\mathbf{CY}+\mathbf{DZ}$ both make sense.
   a. Prove this bilinearity.
   b. Rephrase $\mathrm{cov}(\mathbf{AW}+\mathbf{BX},\mathbf{CY}+\mathbf{DZ})$ in terms of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, $\boldsymbol{\Sigma}_{\mathbf{WY}}$, $\boldsymbol{\Sigma}_{\mathbf{WZ}}$, $\boldsymbol{\Sigma}_{\mathbf{XY}}$ and $\boldsymbol{\Sigma}_{\mathbf{XZ}}$.
2. Let $\mathbf{A}$ be a square matrix with eigendecomposition $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$. Given a real number c ($\neq$ any eigenvalue of $\mathbf{A}$), express the eigendecomposition of $(\mathbf{A} - c\mathbf{I})^{-1}$ in terms of $\mathbf{U}$, Λ, $\mathbf{I}$ and c.
3. Let W be a discrete random variable such that $\Pr(W = 1) = \Pr(W = -1) = 1/2$. Define $Y = WX$ with $X \sim N(0,1)$ and $X \perp\!\!\!\perp W$. Prove the following identities.
   a. $Y \sim N(0,1)$.
   b. X and Y are uncorrelated with each other.
   c. X is not independent of Y.
4. Let $\mathbf{X} = [X_1, X_2, X_3]^\top \sim MVN_3(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with
$$\boldsymbol{\mu} = [6,1,4]^\top, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}.$$
   a. Find the conditional distribution of X_2 given $X_1 = 2$ and $X_3 = 1$.
   b. Find the distribution of random 2-vector $\mathbf{Y} = [3X_1 - 2X_2 + X_3, X_2 - X_3]^\top$.
   c. Find $w_1, w_2 \in \mathbb{R}$ such that $W = w_1X_1 + w_2X_2 + X_3$ is independent of $\mathbf{Y}$. (Hint: don't forget to verify the normality of random 3-vector $[W, \mathbf{Y}^\top]^\top$ after figuring out values of w_1 and w_2.)
5. Consider the `Wolves` dataset from the package `candisc`. The variable `sex` indicates the sex of wolves (`f=female`, `m=male`), while `location` encodes wolves' habitats (`ar=Artic`, `rm=Rocky Mountain`). The combination of `location` and `sex` is exactly `group`. Variables `x1` to `x9` correspond to 9 different skull morphological measurements of wolves, respectively. **We will merely focus on six measurements `x4` to `x9`.**
   a. Perform an appropriate test to compare the mean skull measurements of male and female wolves. Is there any statistical evidence to claim that the the morphology of the skull differs between males and females at 5% level? (Hint: don't forget to include your hypotheses, name of method, value of test statistic, and rejection region/p-value, before coming to the conclusion.)
   b. What are the assumptions that were required to perform the test in part a?

c. Repeat parts a and b only for wolves from the Arctic.
d. Provide plausible explanations (both statistical and subject-matter) about any discrepancy between the full analysis and the subgroup analysis.
e. Use a formal test to check the heteroscedasticity between males and females.

```
library(candisc)
head(Wolves)
```

```
##      group location sex  x1  x2  x3   x4   x5   x6   x7   x8   x9
## rmm1  rm:m       rm   m 126 104 141 81.0 31.8 65.7 50.9 44.0 18.2
## rmm2  rm:m       rm   m 128 111 151 80.4 33.8 69.8 52.7 43.2 18.5
## rmm3  rm:m       rm   m 126 108 152 85.7 34.7 69.1 49.3 45.6 17.9
## rmm4  rm:m       rm   m 125 109 141 83.1 34.0 68.0 48.2 43.8 18.4
## rmm5  rm:m       rm   m 126 107 143 81.9 34.0 66.1 49.0 42.4 17.9
## rmm6  rm:m       rm   m 128 110 143 80.6 33.0 65.0 46.4 40.2 18.2
```