STAT 415 — Midterm 1 (Spring 2020)

Answer: Solutions should be written below each question.

Name: Solutions

Exam rules:

- You have 75 minutes to complete the exam.
- There are 5 Questions.
- You will get a formula sheet.
- You are only allowed to use a pen and pencil.
- Explain your arguments carefully, show all your work, and write out all your steps. Credit will be given for clarity. No credit will be given for a solution without working.
- If you are caught cheating or helping someone to cheat on this exam, you will both receive a grade of zero on the exam and you will be turned in.
- If a problem requires a numerical answer, you may express answer in terms of elementary functions or cdfs (e.g. $\Gamma(r+1)$ or $r!$).

1. Suppose

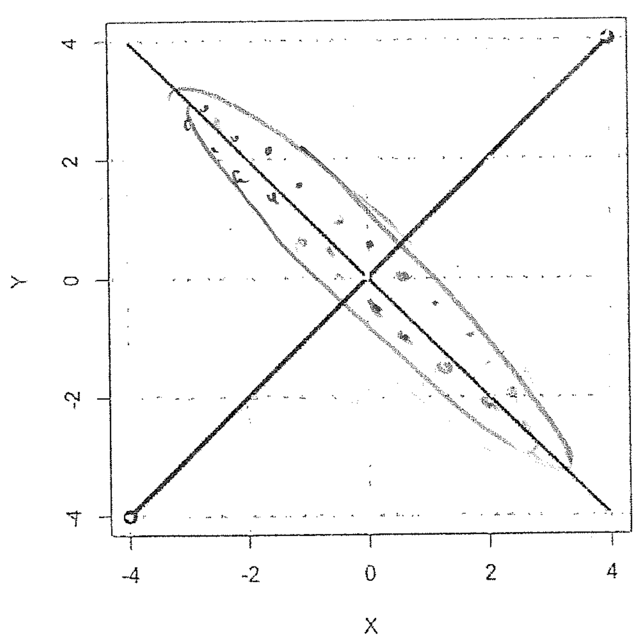
$$\underline{e}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \underline{e}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Define the bivariate random vector

$$\begin{aligned} \underline{X} = \begin{pmatrix} X \\ Y \end{pmatrix} &= 0.2 Z_1 \underline{e}_1 + 2 Z_2 \underline{e}_2 \\ &= \underbrace{0.2}_{\text{Small}} \begin{pmatrix} Z_1 \\ Z_1 \end{pmatrix} + \underbrace{2}_{\text{large}} \begin{pmatrix} Z_2 \\ -Z_2 \end{pmatrix} \end{aligned}$$

where Z_1, Z_2 are iid normal random variables.

(i) Suppose that $\{\underline{X}\}_{i=1}^{n}$ are iid random variables where $\underline{X}_i \sim \underline{X}$. Draw on the plot below 20 or so typical realisations of $\{\underline{X}\}_{i=1}^{n}$ (plotting X against Y), enough to get the general behaviour of $\underline{X}$. [2]



<u>Assume</u>

$Z_1, Z_2 \sim N(0,1)$

$\Rightarrow E(X_1) = 0$

$E(X_2) = 0$

Dots are centered about zero.

(ii) Explain your answer. [2]

- "Small" (0.2 relative to 2) variation on line $\underline{e}_1$
- "large" (2 relative to 0.2) variation on line $\underline{e}_2$.

(*) Combined the variation is largest along $\underline{e}_2$

See Solution to (HW1, Q3)

2. Suppose $\{X_i\}$ and U are independent random variables, where $\mathrm{E}[X_i] = 0$, $\mathrm{var}[X_i] = \sigma_X^2$, $\mathrm{E}[U] = \mu$ and $\mathrm{var}[U] = \sigma_U^2$. Let $Z_i = X_i + U$ and $S_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$.

(i) Calculate $\mathrm{E}[S_n]$. [2]

$$E[S_n] = E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i+U)\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i+U] = \frac{1}{n}\sum_{i=1}^{n}\mu = \mu$$

(ii) Calculate $\mathrm{var}[S_n]$. [4]

$$\mathrm{var}[S_n] = \mathrm{var}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\right] = \frac{1}{n^2}\sum_{i,j=1}^{n}\mathrm{cov}(Z_i, Z_j)$$

Consider two case separately

(a) $i = j$ $\quad \mathrm{cov}(Z_i, Z_j) = \mathrm{var}(Z_i) = \sigma_\varepsilon^2 + \sigma_u^2$

(b) $i \neq j$ $\quad \mathrm{cov}(Z_i, Z_j) = \mathrm{cov}[(X_i+U), (X_j+U)]$
$$= \mathrm{cov}(U, U) = \sigma_u^2.$$

Return to $\mathrm{var}[S_n] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}(Z_i) + \frac{1}{n^2}\sum_{i\neq j}\mathrm{cov}(Z_i, Z_j)$

$$= \frac{n[\sigma_\varepsilon^2 + \sigma_u^2]}{n^2} + \frac{n(n-1)}{n^2}\cdot\sigma_u^2$$

$$= \frac{\sigma_\varepsilon^2}{n} + \sigma_u^2$$

$\boxed{\text{It is important to understand the implications of this identity}}$

(See HW1, Q7)

(*) See Section 1.5.1

(iii) Calculate the mean squared error $\mathrm{E}[S_n - \mu]^2$. What happens to the mean squared error as $n \to \infty$? [2]

Recall $\mathrm{var}(S_n) = E\left(S_n - E[S_n]\right)^2 = E[S_n - \mu]^2$

$$= \frac{\sigma_\varepsilon^2}{n} + \sigma_u^2 \quad \text{(From part ii)}$$

$$\xrightarrow{n\to\infty} \sigma_u^2$$

(iv) Is S_n an asymptotically consistent estimator of μ, either in probability or mean squared error. Give a reason for your answer. [2]

(*) Consistent means estimator converges to population parameter as $n \to \infty$

If $E[S_n - \mu]^2 \longrightarrow 0$ as $n \to \infty$ then the S_n is a consistent estimator of μ [since the mean squared distance between S_n and μ get closer together

But $E[S_n - \mu]^2 \longrightarrow \sigma_u^2 \neq 0$ hence S_n is not a consistent estim of μ.

(3) Suppose that $\{Z_i\}_{i=1}^3$ are iid standard normal random variables ($Z_i \sim N(0,1)$).

(i) Let $Y_1 = Z_1 + Z_2$, $Y_2 = Z_2$ and $Y_3 = Z_2 + Z_3$.
Calculate the variance matrix of $\underline{Y}' = (Y_1, Y_2, Y_3)$. Draw the associated covariance network. [6]

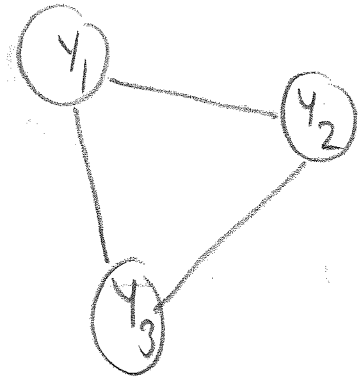
$\mathrm{var}(Y_1) = \mathrm{var}(Z_1) + \mathrm{var}(Z_2)$ (by independence)
$= 2$

$\mathrm{var}(Y_2) = 1$, $\mathrm{var}(Y_3) = 2$

$\mathrm{cov}(Y_1, Y_2) = \mathrm{var}(Z_2) = 1$ $\quad$ $\mathrm{cov}(Y_1, Y_3) = \mathrm{var}(Y_2) = 1$

$\mathrm{cov}(Y_2, Y_3) = 1$

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \mathrm{var}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$



$$\operatorname{var}(Z_2 Z_3) = E[Z_2^2 Z_3^2] - (E[Z_2 Z_3])^2$$
$$= E(Z_2^2)E(Z_3^2) - 0^2 = 1$$

(ii) Let $X_1 = Z_1 + Z_2Z_3$, $X_2 = Z_2$ and $X_3 = Z_2 + Z_3$.

Calculate the variance matrix of $\underline{X}' = (X_1, X_2, X_3)$. Draw the associated covariance network. [5]

$\operatorname{Var}(X_1) = \operatorname{Var}(Z_1) + \operatorname{var}(Z_2Z_3) = 1 + 1 = 2$ (since Z_1 es indepen of Z_2Z_3)

$\operatorname{var}(X_2) = \operatorname{var}(Z_2) = 1$

$\operatorname{var}(X_3) = \operatorname{var}(Z_2) + \operatorname{var}(Z_3) = 2$

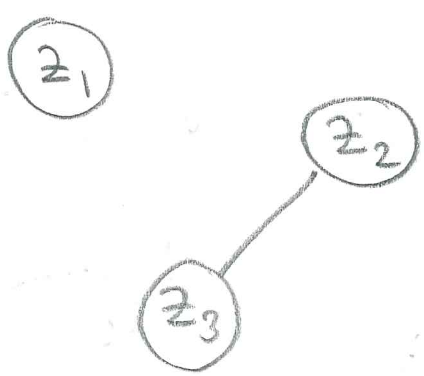
$$\operatorname{cov}(X_1, X_2) = \operatorname{cov}[Z_1 + Z_2Z_3, Z_2] = \operatorname{cov}[Z_2Z_3, Z_2]$$
$$= E(Z_2^2 Z_3) - E(Z_2Z_3)E(Z_2)$$
$$= E(Z_2^2)E(Z_3) - E(Z_2)E(Z_3)E(Z_2)$$
$$= 0$$

$$\operatorname{cov}(X_1, X_3) = \operatorname{cov}[Z_1 + Z_2Z_3, Z_2 + Z_3]$$
$$= \operatorname{cov}[Z_2Z_3, Z_2] + \operatorname{cov}[Z_2Z_3, Z_3] = 0 \text{ (by)}$$

$$\operatorname{cov}(X_2, X_3) = \operatorname{cov}(Z_2, Z_2) = 1$$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \operatorname{var}\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix}$$

Question What happens if $E(Z_1) = E(Z_2) = E(Z_3) = 1$?



(*) Since X_1 is not normal the fact that X_1 is uncorrelated with X_2 and X_3 does not mean that it is independent of X_2 or X_3.

See Section 2.4.2 and HW4, Q3.

4. Suppose that X_1 and X_2 are iid normal random variables with mean μ and variance σ^2. Let $\bar{X} = 2^{-1}(X_1 + X_2)$ and $s_2^2 = (X_1 - \bar{X})^2 + (X_2 - \bar{X})^2$.

(i) Show that $\bar{X}$ and s_2^2 are independent random variables (credit will be given for a coherent argument, state precisely all results that you use). [6]

Hint: Use that $s_2^2 = (X_1 - X_2)^2/2$.

Observe $s_2^2 = \left[\frac{(X_1 - X_2)}{\sqrt{2}}\right]^2 = Y^2$ where $Y = \frac{(X_1 - X_2)}{\sqrt{2}}$

Since $\binom{X_1}{X_2}$ are jointly normal then $\binom{\bar{X}}{Y}$ are jointly normal too.

If we can show that $\bar{X}$ and Y are uncorrelated, normality implies that $\bar{X}$ and Y are independent. This, then, implies $\bar{X}$ and $Y^2 = s_2^2$ are independent. Thus proving the result.

Aim show $\bar{X}$ and Y are uncorrelated (and thus independent)

$$\mathrm{cov}[\bar{X}, Y] = \mathrm{cov}\left[\frac{1}{2}(X_1 + X_2), \frac{1}{\sqrt{2}}(X_1 - X_2)\right]$$

$$= \frac{1}{2^{3/2}} \mathrm{cov}(X_1 + X_2, X_1 - X_2)$$

$$\overset{\text{By independence of } X_1 \text{ and } X_2}{=} \frac{1}{2^{3/2}}\left[\mathrm{cov}(X_1, X_1) - \mathrm{cov}(X_2, X_2)\right]$$

$$= \frac{1}{2^{3/2}}[1 - 1] = 0.$$

Thus proving the result!

(ii) Define the T and Z-transforms: $T_2 = \frac{\sqrt{2}(\bar{X}-\mu)}{s_2}$ and $Z = \frac{\sqrt{2}(\bar{X}-\mu)}{\sigma}$. Calculate

$$P\left(T_2 > \frac{3}{2}Z\right).$$

Give your answer in terms of an appropriate χ^2 distribution. You may use that $s_2^2 \sim \sigma^2\chi_1^2$. [3]

$$P\left[\frac{\sqrt{2}\cancel{(\bar{X}-\mu)}}{S_2} > \frac{3}{2}\cdot\frac{\sqrt{2}\cancel{(\bar{X}-\mu)}}{\sigma}\right]$$

$$= P\left[\frac{1}{S_2} > \frac{3}{2}\cdot\frac{1}{\sigma}\right]$$

$$= P\left[S_2 < \frac{2}{3}\cdot\sigma\right\} = P\left\{S_2^2 \leq \frac{4}{9}\cdot\sigma^2\right.$$

$$S_2^2 = \sigma^2\chi_1^2$$

$$= P\left\{\cancel{\sigma^2}\chi_1^2 \leq \frac{4}{9}\cdot\cancel{\sigma^2}\right\}$$

$$= P\left\{\chi_1^2 \leq \frac{4}{9}\right\}$$

(See HW 5, Q 1)

(*) See HW5, Q2.

5. Suppose that $\{X_i\}$ are iid random variables drawn from an exponential distribution with density $\lambda \exp(-\lambda x)$ where $x \geq 0$.

(i) Let $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. State the limiting distribution of $\bar{X}_n$ (giving the correct mean and variance). [3]

Use the formula $E(X) = \frac{1}{\lambda}$ $E(X^2) = \frac{2}{\lambda^2}$

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an estimator of $E(X) = \frac{1}{\lambda}$

$$\boxed{\begin{aligned} &\mathrm{Var}(X) \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}}$$

By using the CLT we have

$$\sqrt{n}\,(\bar{X} - \underbrace{\tfrac{1}{\lambda}}_{\mu}) \xrightarrow{D} N(0, \underbrace{\tfrac{1}{\lambda^2}}_{\sigma^2})$$

(ii) Obtain the first moment of X_i and the method of moments estimator of λ (based on the first moment)? [3]

Since $E(X) = \frac{1}{\lambda}$, then $\lambda = \frac{1}{E(X)}$

For the moments estimator replace $E(X)$ with its estimator which is $\bar{X}$.

Moments estimator $\boxed{\hat{\lambda}_n = \frac{1}{\bar{X}}}$