# STAT 510, Midterm Exam

March 12, 2020

**Name:** TIANQI WU

**NetID:** tw38

**Please sign the following pledge and read all instructions carefully before starting the exam.**

**Pledge:** I have neither given nor received any unauthorized aid in completing this exam, and I have conducted myself within the guidelines of the University Student Code.

**Signature:** ______________________________

**INSTRUCTIONS:**

- This is a closed-book exam. However, you are allowed to bring an 8.5×11 sheet of notes (two-sides).
- Total time is 80 minutes (08:00 A.M to 09:20 A.M.)
- **Show all work,** clearly and in order, if you want to receive full credit. I reserve the right to take off points if I cannot see how you arrived at your answer (even if your final answer is correct).
- **Answer all the questions in the space provided. You may attach additional sheets if necessary.**
- This test has 4 regular questions and one bonus question (100 + 5 points). It is your responsibility to make sure that you have all of the questions.
- **Good luck!**

| Que. No. | Max Points | Earned Pts. |
|---|---|---|
| 1 | 25 | |
| 2 | 25 | |
| 3 | 25 | |
| 4 | 25 | |
| 5 (bonus) | 5 | |

**TOTAL:** __________

**Question 1.** Let $X_i$ follow the Binomial distribution $\text{Bin}(n_i, p)$ for $i = 1, 2, \ldots, m$. Assume they are independent.

a. (12 points) Find the moment generating function of each $X_i$;

b. (13 points) What is the distribution of their sum $S_m = \sum_{i=1}^{m} X_i = X_1 + \cdots + X_m$?

(a) $E[e^{tx_i}] = \sum_{x_i=0}^{n_i} e^{tx_i} f_{x_i}(x_i)$

$= \sum_{x_i=0}^{n_i} e^{tx_i} \binom{n_i}{x_i} p^{x_i}(1-p)^{n_i - x_i}$

$= \sum_{x_i=0}^{n_i} \binom{n_i}{x_i} (pe^t)^{x_i} (1-p)^{n_i - x_i}$

$= (pe^t + 1 - p)^{n_i}$

(b) $M_{S_m}(t) = E[e^{tS_m}]$

$= E[e^{t(x_1 + x_2 + \cdots + x_m)}]$

Since $x_i$ are indepedent. $= E[e^{tx_1}] \cdot E[e^{tx_2}] \cdots E[e^{tx_m}]$

$= (pe^t + 1 - p)^{n_1} \cdots (pe^t + 1 - p)^{n_m}$

$= (pe^t + 1 - p)^{\sum_{i=1}^{m} n_i}$

Hence, $S_m \sim \text{Bin}(\sum_{i=1}^{M} n_i, p)$

**Question 2.** Let $\mathbf{X} = (X_1, X_2, X_3)^T$ be $N(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu}^{\mathrm{T}} = (\mathbf{2}, \mathbf{-3}, \mathbf{1})$ and $\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}$.

a. (12 points) Find the distribution of $3X_1 - 2X_2 + X_3$.

b. (13 points) Find a vector $\boldsymbol{a} \in \mathbb{R}^2$ such that $X_2$ is independent of $X_2 - \boldsymbol{a}^T \begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$.

(a). Since $X = \mu + BZ$ where $Z \overset{iid}{\sim} N(0,1)$. $\Sigma = BB'$

$DX = D(\mu + BZ) = D\mu + (DB)Z$

Hence $DX \sim N(D\mu, DBB'D')$

$DX \sim N(D\mu, D\Sigma D')$, let $D = (3, -2, 1)$

$D\mu = (3, -2, 1) \cdot \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = 13$

$D\Sigma D' = (3, -2, 1) \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = 9$

Hence, $3X_1 - 2X_2 + X_3 \sim \boxed{N(13, 9)}$

(b). To let $X_2$ independent of $X_2 - (a_1, a_2)\begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$

$\Rightarrow Cov(X_2, X_2 - a_1 X_1 - a_2 X_3) = 0$

$Cov(X_2, X_2) - a_1 Cov(X_2, X_1) - a_2 Cov(X_2, X_3) = 0$

$3 - a_1 - 2a_2 = 0$

$a_1 + 2a_2 = 3$

let $a_1 = k$ $\quad a_2 = \frac{3-k}{2}$

$\boxed{a = \begin{pmatrix} k \\ \frac{3-k}{2} \end{pmatrix}}$, $k \in \mathbb{R}$

**Question 3.** (25 points) Let $X_i$, $i = 1, 2, \ldots, n$, be iid random variables following $N(\mu, \sigma^2)$. Show that the sample mean $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ are independent.

First, consider joint distribution $(\bar{X}_n, U)$ where $U = \sum_{i=1}^n (X_i - \bar{X}_n)^2$

Since $X \sim N(\mu 1_n, \sigma^2 I_n)$, $\bar{X}_n = \frac{1}{n} 1_n' X$

$$\begin{pmatrix} X_1 - \bar{X}_n \\ \vdots \\ X_n - \bar{X}_n \end{pmatrix} = X - 1_n \bar{X}_n = I_n X - \frac{1}{n} 1_n 1_n' X = H_n X, \text{ where } H_n = I_n - \frac{1}{n} 1_n 1_n'$$

Stack them: $$\begin{pmatrix} \bar{X}_n \\ X_1 - \bar{X}_n \\ \vdots \\ X_n - \bar{X}_n \end{pmatrix} = \begin{pmatrix} \frac{1}{n} 1_n' \\ H_n \end{pmatrix} X$$

Since it is a linear transformation of multivarite normal.

$$\mathrm{Cov}\left[\begin{pmatrix} \bar{X}_n \\ X_1 - \bar{X}_n \\ \vdots \\ X_n - \bar{X}_n \end{pmatrix}\right] = \begin{pmatrix} \frac{1}{n} 1_n' \\ H_n \end{pmatrix} \sigma^2 I_n \begin{pmatrix} \frac{1}{n} 1_n' \\ H_n \end{pmatrix}'$$

$$= \sigma^2 \begin{pmatrix} \frac{1}{n^2} 1_n' 1_n & \frac{1}{n} 1_n' H_n \\ \frac{1}{n} H_n 1_n & H_n H_n \end{pmatrix}$$

$$= \sigma^2 \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & H_n \end{pmatrix}$$

Since $1_n' H_n = 1_n' (I_n - \frac{1}{n} 1_n 1_n') = 1_n' - \frac{1}{n} \cdot n \cdot 1_n' = 0$

$H_n 1_n = (I_n - \frac{1}{n} 1_n 1_n') 1_n = 1_n - \frac{1}{n} \cdot n\, 1_n = 0$

Hence, $\bar{X}_n$ and $H_n X$ are independent, $U$ only depends on $H_n X$.

$\bar{X}_n$ and $U = \|H_n X\|^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2$ are independent.

$\bar{X}_n$ and $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ are independent.

扫描全能王 创建

**Question 4.** (25 points) Let $X_1, X_2, \ldots, X_n$ be a sample from the inverse Gaussian distribution whose pdf is

$$f(x|\mu,\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\{-\lambda(x-\mu)^2/(2\mu^2 x)\}, \quad x > 0,$$

where $\lambda \in (0,\infty)$ and $\mu \in (0,\infty)$. Find the maximum likelihood estimators (MLE) of $\mu$ and $\lambda$.

$$L(\mu,\lambda) = \prod_{i=1}^{n}\left(\frac{\lambda}{2\pi x_i^3}\right)^{1/2} \exp\left(-\sum_{i=1}^{n} -\lambda(x_i-\mu)^2/(2\mu^2 x_i)\right)$$

$$l(\mu,\lambda) = C + \frac{n}{2}\ln(\lambda) - \sum_{i=1}^{n}\frac{\lambda(x_i-\mu)^2}{2\mu^2 x_i} = C + \frac{n}{2}\ln(\lambda) - \sum_{i=1}^{n}\frac{\lambda x_i^2 - 2\lambda x_i\mu + \lambda\mu^2}{2\mu^2 x_i}$$

$$\frac{dl}{d\mu} = \sum_{i=1}^{n}\frac{\lambda x_i}{\mu^3} - \frac{\lambda}{\mu^2} = \frac{n\lambda}{\mu^3}(\bar{x}-\mu) = 0 \Rightarrow \boxed{\hat{\mu} = \bar{x}}$$

$$\frac{dl}{d\lambda} = \frac{n}{2\lambda} - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\mu^2 x_i} \text{, plug in } \hat{\mu} = \bar{x} \Rightarrow$$

$$= \frac{n}{2\lambda} - \sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{2\bar{x}^2 x_i}$$

$$= \frac{n}{2\lambda} - \sum_{i=1}^{n}\frac{(x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{2\bar{x}^2 x_i}$$

$$= \frac{n}{2\lambda} - \sum_{i=1}^{n}\left(\frac{x_i}{2\bar{x}^2} - \frac{1}{\bar{x}} + \frac{1}{2x_i}\right)$$

$$= \frac{n}{2\lambda} - \left(\frac{n\bar{x}}{2\bar{x}^2} - \frac{n}{\bar{x}} + \frac{1}{2}\sum_{i=1}^{n}\frac{1}{x_i}\right)$$

$$= \frac{n}{2}\left(\frac{1}{\lambda} - \left(\frac{1}{\bar{x}} - \frac{2}{\bar{x}} + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}\right)\right) = 0 \Rightarrow$$

$$\boxed{\begin{aligned} \frac{1}{\hat{\lambda}} &= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i} - \frac{1}{\bar{x}} \\ \hat{\lambda} &= \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i} - \frac{1}{\bar{x}}} \end{aligned}}$$

**Question 5. (bonus)** (5 points) Let $X_1, \ldots, X_n$ be iid continuous random variables with cumulative distribution (cdf) $F$ and probability density function (pdf) $f$. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the order statistics. Find the joint pdf of $(X_{(1)}, X_{(n)})$.

Using probability transform approach,

$X \overset{D}{=} (F^{-1}(U_1) \cdots F^{-1}(U_n))$, where $U_i$' $\overset{iid}{\sim}$ Uniform$(0,1)$.

Since $F$ is cdf, $F^{-1}$ is increasing:

$(X_{(1)} \cdots X_{(n)}) \overset{D}{=} (F^{-1}(U_{(1)}) \cdots F^{-1}(U_{(n)}))$

$X_{(k)} \overset{D}{=} F^{-1}(U_{(k)})$, $U_{(k)} \sim Beta(k, n-k+1)$

$U_{(1)} \sim Beta(1, n)$ $U_{(n)} \sim Beta(n, 1)$

Pdf of $(U_{(1)}, U_{(n)})$ is $n(n-1)(U_{(n)} - U_{(1)})^{n-2}$

$$J_{X_{(1)}, X_{(n)}}(x_1, x_n) = \begin{vmatrix} \frac{dF(X_{(1)})}{dx_1} & 0 \\ 0 & \frac{dF(X_{(n)})}{dx_n} \end{vmatrix} = f(x_1) f(x_n)$$

Pdf of $(X_{(1)}, X_{(n)}) = n(n-1)(F(X_{(n)}) - F(X_{(1)}))^{n-2} \cdot f(x_1) f(x_n)$