Qualifying Exam Solutions: Theoretical Statistics

1.
(a) For the first sampling plan, the expectation of any statistic $W(X_1, X_2, \ldots, X_n)$ is a polynomial of θ of degree less than $n+1$. Hence $\tau(\theta)$ cannot have an unbiased estimator from this plan. For the second plan, let Y be the number of X_i's observed. Then

$$P(Y = y) = \theta(1-\theta)^{y-1}, y = 1, 2, \ldots;$$

and $E_\theta(Y) = 1/\theta$. Therefore, Y is an unbiased estimator of θ and the second sampling plan should be employed.
(b) Observe X_i's till you get two X_i's equal to 1. Let Y_1 be the the number of X_i's till the first 1, and Y_2 be the number of X_i's between the first and second ones. Then Y_1 and Y_2 are iid, and $E_\theta(Y_i) = 1/\theta$ implies that $E_\theta(Y_1 Y_2) = 1/\theta^2$.

2.
(a) The joint pmf is $(1/N)^n \mathrm{I}_{X_{(n)} \leq N} \mathrm{I}_{X_{(1)} \geq 1}$ where X_i's are in $\{1, 2, \ldots, N, N+1, \ldots\}$ and $X_{(r)}$ is the rth order statistic. By the Factorization Theorem, $X_{(n)}$ is sufficient. We show it is complete.

$$P(X_{(n)} \leq k \mid N) = P(X_1 \leq k \mid N) \cdots P(X_n \leq k \mid N) = (\frac{k}{N})^n$$

if $1 \leq k \leq N$. And

$$P(X_{(n)} = k \mid N) = P(X_{(n)} \leq k \mid N) - P(X_{(n)} \leq k-1 \mid N) = (\frac{k}{N})^n - (\frac{k-1}{n})^n.$$

Let $h(X_{(n)})$ have $E_N(h(X_{(n)}) = o$ for any $N = 1, 2, \ldots$. Putting $N = 1$, $h(1) = 0$; $N = 2$, $h(1)P(X_{(n)} = 1) + h(2)P(X_{(n)} = 2) = h(2) = 0$ since $P(X_{(n)} = 2 \mid N) > 0$ for $N \geq 2$. Proceeding this way and by induction, $h = 0$ identically. Therefore, X_n is complete.
(b)An unbiased estimator of N is $2X_1 - 1$. Hence the BUE of N is $E(2X_1 - 1 \mid X_{(n)}) = 2E(X_1 \mid X_{(n)}) - 1$.
(Suggestion from Professor Ghosh: it is already acceptable if a student stops here)
Incidentally,

$$E(X_1 \mid X_{(n)} = x_{(n)}) = x_{(n)} P(X_1 = x_{(n)} \mid X_{(n)} = x_{(n)}) + \sum_{1 \leq x < x_{(n)}} x P(X_1 = x \mid X_n = x_{(n)})$$

and

$$P(X_1 = x_{(n)} \mid X_{(n)} = x_{(n)}) = \sum_{k=1}^{n} (\frac{1}{N})^k \begin{pmatrix} n-1 \\ k-1 \end{pmatrix} \left(\frac{x_{(n)} - 1}{N} \right)^{n-k}$$

and

$$P(X_1 = x < x_{(n)} \mid X_{(n)} = x_{(n)}) = \frac{1}{N} \sum_{k=1}^{n-1} (\frac{1}{N})^k \begin{pmatrix} n-1 \\ k \end{pmatrix} \left(\frac{x_{(n)} - 1}{N} \right)^{n-k-1}$$

from which $E(X_1 \mid X_{(n)} = x_{(n)})$ can be found.

3.
(a) It can be shown or known that T is the complete sufficient statistic of λ. If we can find f such that $E[f(T)] = e^{a\lambda}$, then $f(T)$ is the UMVUE. Consider $f(u) = b^u$. Then

$$E(b^T) = \sum_{k=0}^{\infty} b^k \frac{(n\lambda)^k}{k!} e^{-n\lambda} = e^{-n\lambda} \sum_{k=0}^{\infty} \frac{(bn\lambda)^k}{k!} = e^{-n\lambda + bn\lambda}.$$

Let $b = 1 + a/n$. We have $E[(1 + a/n)^T] = e^{a\lambda}$. Therefore, $(1 + a/n)^T$ is the UMVUE of $e^{a\lambda}$.
(b) Note that

$$E(2^{X_1}) = \sum_{k=0}^{\infty} 2^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda}.$$

So 2^{X_1} is an unbiased estimator of e^{λ}. Because T is complete and sufficient, $E(2^{X_1}|T)$ must be the UMVUE of e^{λ}. Since UMVUE is unique, using the result in (a), we have

$$E(2^{X_1}|T) = (1 + \frac{1}{n})^T.$$

4.
(a) Use $\pi_{a,b} = \text{Beta}(a, b)$ as the prior distribution. Then the posterior distribution of p is $\text{Beta}(a + x, n - x + b)$, and the Bayes estimator of p is

$$\hat{p}_{a,b} = \frac{a + x}{n + a + b} = \frac{n}{n + a + b}\hat{p} + \frac{a}{n + a + b}.$$

When $\alpha > 0$, $\beta < 0$ and $\alpha + \beta < 1$, we can let $a = n\beta/\alpha$ and $b = n(1 - \alpha - \beta)/\alpha$, both of which are positive, so that $\delta(x)$ is a Bayes estimator and thus is admissible.
(**More sophisticated proofs are possible, for example, by using Karlin's Theorem, which leads to stronger conclusions**)
(b) The risk function of $\delta(x)$ is

$$\begin{aligned} R(\delta, p) =& E_p[(\alpha\hat{p} + \beta - p)^2] \\ =& E_p(\alpha\hat{p} - \alpha p)^2 + [\beta - (1 - \alpha)p]^2 \\ =& \frac{\alpha^2 p(1 - p)}{n} + [\beta - (1 - \alpha)p]^2 \\ =& [(1 - \alpha)^2 - \frac{\alpha^2}{n}]p^2 + [\frac{\alpha^2}{n} - 2\beta(1 - \alpha)]p + \beta^2. \end{aligned}$$

Setting

$$(1 - \alpha)^2 - \frac{\alpha^2}{n} = \frac{\alpha^2}{n} - 2\beta(1 - \alpha) = 0,$$

that is, setting

$$\alpha = \frac{\sqrt{n}}{1+\sqrt{n}}, \beta = \frac{1}{2(1+\sqrt{n})}$$

, $R(\delta, p)$ becomes a constant. Therefore, the corresponding estimator

$$\delta(X) = \frac{\sqrt{n}}{1+\sqrt{n}} X + \frac{1}{2(1+\sqrt{n})}$$

is minimax. As a matter of factor, this $\delta(X)$ is the Bayes estimator corresponding to the prior distribution Beta($\sqrt{n}/2, \sqrt{n}/2$).

5.
(a) Apply the Neyman-Pearson Lemma, the rejection region of a size α UMP test includes $x_1, x_2, \ldots, x_n$ such that

$$\begin{aligned}\frac{f(x_1,\ldots,x_n|H_a)}{f(x_1,\ldots,x_n|H_0)} &= \frac{(2\pi\sigma^2)^{-n/2}\exp(-\frac{1}{2\sigma^2}\sum(x_i-\theta_{i0})^2)}{(2\pi\sigma^2)^{-n/2}\exp(-\frac{1}{2\sigma^2}\sum x_i^2)} \\ &= \exp(\frac{1}{\sigma^2}\sum x_i\theta_{i0})\exp(-\frac{1}{2\sigma^2}\sum \theta_{i0}^2) > k,\end{aligned}$$

or equivalently, $\sum x_i\theta_{i0} > k'$, with $P(\sum X_i\theta_{i0} > k'|H_0) = \alpha$. Since X_i's are independent of each other, the distribution of $\sum X_i\theta_{i0}$ is $N(0, \sigma^2\sum\theta_{i0}^2)$ under H_0. Therefore, we can choose $k' = \sigma z_\alpha\sqrt{\sum\theta_{i0}^2}$, and the rejection region of the size α UMP test is

$$\Big\{(x_1,\ldots,x_n) \Bigm| \sum x_i\theta_{i0} > \sigma z_\alpha\sqrt{\sum\theta_{i0}^2}\Big\}.$$

(b). The MLE of θ_i is $\hat{\theta}_i = X_i$ for $1 \le i \le n$. So the LRT statistic is

$$\lambda = \frac{(2\pi\sigma^2)^{-n/2}\exp(-\frac{1}{2\sigma^2}\sum x_i^2)}{(2\pi\sigma^2)^{-n/2}\exp(-\frac{1}{2\sigma^2}\sum(x_i-\hat{\theta}_i)^2)} = \exp(-\frac{1}{2\sigma^2}\sum x_i^2)$$

The reject region is $\lambda < c$ for some constant c such that $0 < c < 1$, or equivalently $\sum X_i^2 > c'$ for some constant c'. Let $T = \sum X_i^2$. Under H_0, $T/\sigma^2 \sim \chi_n^2$. So H_0 is rejected when $T = \sum_{i=1}^n X_i^2 > \sigma^2\chi^2_{n,\alpha}$. Using normal approximation, the rejection region is $\sum x_i^2 > n\sigma^2 + \sigma\sqrt{2n}z_\alpha$.
(c). For the test obtained in (a), under H_a, $\sum X_i\theta_{i0} = n^{-1/3}\sum X_i$ follows a normal distribution with mean $n^{1/3}$ and variance $n^{1/3}\sigma^2$. So its power is

$$1 - \Phi\Big(\frac{\sigma z_\alpha n^{1/6} - n^{1/3}}{\sigma n^{1/6}}\Big) = 1 - \Phi(z_\alpha - \sigma^{-1}n^{1/6}) \to 1 \text{ as } n \to \infty$$

where Φ is the cdf of the standard normal distribution.

For the test obtained in (b), under H_a, $\sum X_i^2$ approximately follows a normal distribution with mean $n\sigma^2 + n^{1/3}$ and variance $2n\sigma^2 + 4n^{1/3}$. So its power is, as $n \to \infty$,

$$1 - \Phi\Big(\frac{n\sigma^2 + \sigma\sqrt{2n}z_\alpha - (n\sigma^2 + n^{1/3})}{\sqrt{2n\sigma^2 + 4n^{1/3}}}\Big) = 1 - \Phi\Big(\frac{z_\alpha - n^{-1/6}/(\sqrt{2}\sigma)}{\sqrt{1 + 2n^{-2/3}\sigma^{-2}}}\Big) \approx \alpha.$$

6.

(a)

Method 1:

Let $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ and

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ -1 & \frac{1}{2} \\ -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then, the null hypothesis is

$$H_0 : A^t\mu = 0.$$

Since the columns of A are not orthogonal, we need to write A as

$$\tilde{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2\sqrt{5}} \\ -\frac{1}{2} & \frac{3}{2\sqrt{5}} \\ -\frac{1}{2} & -\frac{3}{2\sqrt{5}} \\ \frac{1}{2} & -\frac{1}{2\sqrt{5}}. \end{pmatrix}$$

Then, we can write $H_0 : \tilde{A}^t\mu = 0$. Let $\bar{y} = (\bar{y}_1, \cdots, \bar{y}_4)$. Note that the rank of $\tilde{A}(\tilde{A}^t\tilde{A})^{-1}\tilde{A}^t$ is 2 and $\bar{y} \sim N(\mu, \frac{\sigma^2}{5}I)$. Then under the null hypothesis, we have

$$\begin{aligned} \bar{y}^t\tilde{A}(\tilde{A}^t\tilde{A})^{-1}\tilde{A}^t\bar{y} =& \frac{1}{4}(\bar{y}_{1\cdot} - \bar{y}_{2\cdot} - \bar{y}_{3\cdot} + \bar{y}_{4\cdot})^2 + \frac{1}{20}(\bar{y}_{1\cdot} + 3\bar{y}_{2\cdot} - 3\bar{y}_{3\cdot} - \bar{y}_{4\cdot})^2 \\ \sim& \frac{\sigma^2}{5}\chi^2_2. \end{aligned}$$

Let

$$F^* = \frac{1}{MSE}[\frac{5}{8}(\bar{y}_{1\cdot} - \bar{y}_{2\cdot} - \bar{y}_{3\cdot} + \bar{y}_{4\cdot})^2 + \frac{1}{8}(\bar{y}_{1\cdot} + 3\bar{y}_{2\cdot} - 3\bar{y}_{3\cdot} - \bar{y}_{4\cdot})^2]$$

where

$$MSE = \frac{1}{15}\sum_{i=1}^{4}\sum_{j=1}^{5}(y_{i\cdot} - y_{\cdot\cdot})^2.$$

Then, under the null hypothesis

$$F^* \sim F_{2,15}$$

and H_0 is rejected if $F^* \geq F_{\alpha,2,15}$, where $F_{\alpha,2,15}$ is the upper α quantile of $F_{2,15}$ distribution.
Method 2:
Directly use the result that the F statistic is

$$F = \frac{\hat{\mu}' A' [A(X'X)^{-1}A']^{-1} A\hat{\mu}/2}{\text{MSE}}$$

where X is the design matrix and $\hat{\mu} = (\bar{y}_{1.}, \bar{y}_{2.}, \bar{y}_{3.}, \bar{y}_{4.})'$. In fact,

$$A'[A(X'X)^{-1}A']^{-1}A = \begin{pmatrix} 3/2 & -1/2 & -2 & 1 \\ -1/2 & 7/2 & -1 & -2 \\ -2 & -1 & 7/2 & -1/2 \\ 1 & -2 & -1/2 & 3/2 \end{pmatrix}$$

(b)
Method 1
Directly use the result that

$$\hat{\mu}_{H_0} = \hat{\mu} + (X'X)^{-1}A'[A(X'X)^{-1}A']^{-1}A\hat{\mu},$$

that is,

$$\begin{pmatrix} \hat{\mu}_{1,h_0} \\ \hat{\mu}_{2,h_0} \\ \hat{\mu}_{3,h_0} \\ \hat{\mu}_{4,h_0} \end{pmatrix} = \begin{pmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \bar{y}_{3.} \\ \bar{y}_{4.} \end{pmatrix} + \frac{1}{5} \begin{pmatrix} 3/2 & -1/2 & -2 & 1 \\ -1/2 & 7/2 & -1 & -2 \\ -2 & -1 & 7/2 & -1/2 \\ 1 & -2 & -1/2 & 3/2 \end{pmatrix} \begin{pmatrix} \bar{y}_{1.} \\ \bar{y}_{2.} \\ \bar{y}_{3.} \\ \bar{y}_{4.} \end{pmatrix}$$

Other methods are also acceptable, including using lagrange multipliers, orthogonal projections, etc

7.
(a) The density of $\vec{X}_i$ is

$$\begin{aligned} f(x_1, x_2) =& \frac{1}{\sqrt{1-\rho^2}(2\pi)\sigma_1\sigma_2} \\ & \times \exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}]\}. \end{aligned}$$

The conditional density of X_{i2} given X_{i1} is

$$f_{X_2|X_1}(x_2|x_1) = \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_2} \exp\{-\frac{1}{2(1-\rho^2)\sigma_2^2}[(x_2-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)]^2\}$$

Therefore, $\tau^2 = (1-\rho^2)\sigma_2^2$.

Then, we have the moment estimator as $\hat{\mu}_1 = \bar{X}_1$, $\hat{\mu}_2 = \bar{X}_2$, $\hat{\sigma}_1^2 = S_1^2$, $\hat{\sigma}_2^2 = S_2^2$, $\rho = S_{12}/(S_1 S_2)$, and $\hat{\tau}^2 = (1-\hat{\rho}^2)\hat{\sigma}_2^2$.
(b) Clearly, $V(X_{i1}) = 2\sigma_1^4$, $V(X_{i2}) = 2\sigma_2^4$. Note that

$$X_{i2}|X_{i1} \sim N(\rho\frac{\sigma_2}{\sigma_1}X_{i1}, (1-\rho^2)\sigma_2^2).$$

For the other terms, we need

$$\begin{aligned} E(X_{i1}^2 X_{i2}^2) =& E[X_{i1}^2 E(X_{i2}^2|X_{i1})] \\ =& E\{X_{i1}^2[(1-\rho^2)\sigma_2^2 + \rho^2\frac{\sigma_2^2}{\sigma_1^2}X_{i1}^2]\} \\ =& (1-\rho^2)\sigma_2^2\sigma_1^2 + 3\rho^2\frac{\sigma_2^2}{\sigma_1^2}\sigma_1^4 \\ =& (1+2\rho^2)\sigma_1^2\sigma_2^2. \end{aligned}$$

We also need

$$E(X_{i1}^3 X_{i2}) = E[X_{i1}^3 E(X_{i2}|X_{i1})] = \frac{\rho\sigma_2}{\sigma_1}E(X_{i1}^3) = \rho\sigma_1^3\sigma_2$$

and

$$E(X_{i2}^3 X_{i1}) = 3\rho\sigma_1\sigma_2^3.$$

Therefore, we have

$$\begin{aligned} V(X_{i1}X_{i2}) =& (1+\rho^2)\sigma_1^2\sigma_2^2 \\ Cov(X_{i1}^2, X_{i2}^2) =& 2\rho^2\sigma_1^2\sigma_2^2 \\ Cov(X_{i1}^2, X_{i1}X_{i2}) =& 2\rho\sigma_1^3\sigma_2 \\ Cov(X_{i2}^2, X_{i1}X_{i2}) =& 2\rho\sigma_1\sigma_2^3 \end{aligned}$$

Therefore, we have the covariance matrix of $(X_{i1}^2, X_{i2}^2, X_{i1}X_{i2})$ equal to

$$\Sigma = \begin{pmatrix} 2\sigma_1^4 & 2\rho^2\sigma_1^2\sigma_2^2 & 0 \\ 2\rho^2\sigma_1^2\sigma_2^2 & 2\sigma_2^4 & 0 \\ 0 & 0 & (1+\rho^2)\sigma_1^2\sigma_2^2 \end{pmatrix}$$

(c) To compute the limiting distribution of ρ_M. We define $g(x,y,z) = z/\sqrt{xy}$. Then,

$$\begin{aligned} \frac{\partial g(x,y,z)}{\partial x} =& -\frac{z}{2\sqrt{x^3 y}} \\ \frac{\partial g(x,y,z)}{\partial y} =& -\frac{z}{2\sqrt{xy^3}} \\ \frac{\partial g(x,y,z)}{\partial z} =& \frac{1}{\sqrt{xy}}. \end{aligned}$$

Let $x = \sigma_1^2$, $y = \sigma_2^2$ and $z = \rho\sigma_1\sigma_2$. We have

$$\frac{\partial g(x,y,z)}{\partial x} = -\frac{\rho}{2\sigma_1^2}$$

$$\frac{\partial g(x,y,z)}{\partial y} = -\frac{\rho}{2\sigma_2^2}$$

and

$$\frac{\partial g(x,y,z)}{\partial z} = \frac{1}{\sigma_1\sigma_2}.$$

Then, we have the limiting variance of $\sqrt{n}(\hat{\rho} - \rho)$ as

$$\begin{pmatrix} -\frac{\rho}{2\sigma_1^2} & -\frac{\rho}{2\sigma_2^2} & \frac{1}{\sigma_1\sigma_2} \end{pmatrix} \begin{pmatrix} 2\sigma_1^4 & 2\rho^2\sigma_1^2\sigma_2^2 & 2\rho\sigma_1^3\sigma_2 \\ 2\rho^2\sigma_1^2\sigma_2^2 & 2\sigma_2^4 & 2\rho\sigma_1\sigma_2^3 \\ 2\rho\sigma_1^3\sigma_2 & 2\rho\sigma_1\sigma_2^3 & (1+\rho^2)\sigma_1^2\sigma_2^2 \end{pmatrix} \begin{pmatrix} -\frac{\rho}{2\sigma_1^2} \\ -\frac{\rho}{2\sigma_2^2} \\ \frac{1}{\sigma_1\sigma_2} \end{pmatrix} = (1-\rho^2)^2.$$

Therefore, we have

$$\sqrt{n}(\hat{\rho} - \rho) \sim^{approx} N(0, (1-\rho^2)^2).$$

To compute the limiting distribution of $\hat{\tau}^2$, we define $g(x,y,z) = y - z^2/x$. Then, we have

$$\frac{\partial g(x,y,z)}{\partial x} = \frac{z^2}{x^2}$$

$$\frac{\partial g(x,y,z)}{\partial y} = 1$$

$$\frac{\partial g(x,y,z)}{\partial z} = -\frac{2z}{x}$$

Let $x = \sigma_1^2$, $y = \sigma_2^2$ and $z = \rho\sigma_1\sigma_2$. We have

$$\frac{\partial g(x,y,z)}{\partial x} = \frac{\rho^2\sigma_2^2}{\sigma_1^2}$$

$$\frac{\partial g(x,y,z)}{\partial y} = 1$$

and

$$\frac{\partial g(x,y,z)}{\partial z} = -\frac{2\rho\sigma_2}{\sigma_1}$$

Then, we have

$$\begin{aligned} &\begin{pmatrix} \rho^2\frac{\sigma_2^2}{\sigma_1^2} & 1 & -\frac{2\rho\sigma_2}{\sigma_1} \end{pmatrix} \begin{pmatrix} 2\sigma_1^4 & 2\rho^2\sigma_1^2\sigma_2^2 & 2\rho\sigma_1^3\sigma_2 \\ 2\rho^2\sigma_1^2\sigma_2^2 & 2\sigma_2^4 & 2\rho\sigma_1\sigma_2^3 \\ 2\rho\sigma_1^3\sigma_2 & 2\rho\sigma_1\sigma_2^3 & (1+\rho^2)\sigma_1^2\sigma_2^2 \end{pmatrix} \begin{pmatrix} \frac{\rho^2\sigma_2^2}{\sigma_1^2} \\ 1 \\ -\frac{2\rho\sigma_2}{\sigma_1} \end{pmatrix} \\ =&2(1-\rho^2)^2\sigma_2^4. \end{aligned}$$

Thus, we have

$$\sqrt{n}(\hat{\tau}^2 - \hat{\tau}) \sim^{approx} N(0, 2(1-\rho^2)^2\sigma_2^4).$$