SOLUTIONS TO
STATISTICAL METHODS EXAM

Solutions to Question 1 This question explores the use of moment generating function.

(i) Let $X \sim \chi^2_a$. The moment generating function of X is:

$$\begin{aligned} M_X(t) &= \frac{1}{2^{a/2}\Gamma(a/2)} \int_0^\infty x^{a/2-1} \exp\left(tx - \frac{x}{2}\right) dx \\ &= \frac{1}{2^{a/2}\Gamma(a/2)} \int_0^\infty x^{a/2-1} \exp\left\{-\left(\frac{1}{2} - t\right) x\right\} dx \\ &= \frac{1}{2^{a/2}\Gamma(a/2)(1/2-t)^{a/2}} \int_0^\infty y^{a/2-1} \exp(-y) dy \\ &= \frac{1}{2^{a/2}\Gamma(a/2)(1/2-t)^{a/2}} \Gamma(a/2) \\ &= \frac{1}{(1-2t)^{a/2}}, \end{aligned}$$

where we have made the substitution that $y = (1/2 - t)x$.

(ii) The first four derivatives of $M_X(t)$ are

$$\begin{aligned} M_X^{'}(t) &= a(1-2t)^{-a/2-1}, \\ M_X^{''}(t) &= a(a+2)(1-2t)^{-a/2-2}, \\ M_X^{'''}(t) &= a(a+2)(a+4)(1-2t)^{-a/2-3}, \\ M_X^{''''}(t) &= a(a+2)(a+4)(a+6)(1-2t)^{-a/2-4}. \end{aligned}$$

So, $E(X) = a$, $E(X^2) = a(a+2)$, $E(X^3) = a(a+2)(a+4)$ and $E(X^4) = a(a+2)(a+4)(a+6)$.

(iii) The moment generating function of $S = X_1 + X_2$ is:

$$\begin{aligned} M_S(t) &= E\left[\exp(tX_1 + tX_2)\right] \\ &= E\left[\exp(tX_1)\right] E\left[\exp(tX_2)\right] \\ &= (1-2t)^{-a/2}(1-2t)^{-b/2} \\ &= (1-2t)^{-(a+b)/2}. \end{aligned}$$

(iv) We have $E(S) = E(X_1+X_2) = E(X_1)+E(X_2) = a+b$ and $Var(S) = Var(X_1+X_2) = Var(X_1) + Var(X_2) = 2a + 2b$.

(v) Since

$$M_S(t) = (1-2t)^{-(a+b)/2},$$

it follows that S has a chi-square distribution with $a+b$ degrees of freedom.

Solutions to Question 2 Suppose $\widehat{\theta}$ is an estimator of θ.

(i) $\widehat{\theta}$ is an unbiased estimator of θ if $E(\widehat{\theta}) = \theta$.

(ii) $\widehat{\theta}$ is an asymptotically unbiased estimator of θ if $\lim_{n\to\infty} E(\widehat{\theta}) = \theta$.

(iii) the bias of $\widehat{\theta}$ is $E(\widehat{\theta}) - \theta$.

(iv) the mean squared error of $\widehat{\theta}$ is $E(\widehat{\theta} - \theta)^2$.

(v) $\widehat{\theta}$ is a consistent estimator of θ if $\lim_{n\to\infty} E(\widehat{\theta} - \theta)^2 = 0$, where n is the size of the sample used to calculate $\widehat{\theta}$.

UP TO THIS BOOK WORK.

Let X and Y be uncorrelated random variables. Suppose that X has mean 2θ and variance 4. Suppose that Y has mean θ and variance 2. The parameter θ is unknown.

(i) The biases and mean squared errors of $\widehat{\theta}_1 = (1/4)X + (1/2)Y$ and $\widehat{\theta}_2 = X - Y$ are:

$$
\begin{aligned}
Bias\left(\widehat{\theta}_1\right) &= E\left(\widehat{\theta}_1\right) - \theta \\
&= E\left(\frac{X}{4} + \frac{Y}{2}\right) - \theta \\
&= \frac{E(X)}{4} + \frac{E(Y)}{2} - \theta \\
&= \frac{2\theta}{4} + \frac{\theta}{2} - \theta \\
&= 0,
\end{aligned}
$$

$$
\begin{aligned}
Bias\left(\widehat{\theta}_2\right) &= E\left(\widehat{\theta}_2\right) - \theta \\
&= E\left(X - Y\right) - \theta \\
&= E(X) - E(Y) - \theta \\
&= 2\theta - \theta - \theta \\
&= 0,
\end{aligned}
$$

$$
\begin{aligned}
MSE\left(\widehat{\theta}_1\right) &= Var\left(\widehat{\theta}_1\right) \\
&= Var\left(\frac{X}{4} + \frac{Y}{2}\right) \\
&= \frac{Var(X)}{16} + \frac{Var(Y)}{4} \\
&= \frac{4}{16} + \frac{2}{4} \\
&= \frac{3}{4},
\end{aligned}
$$

and

$$
\begin{aligned}
MSE\left(\widehat{\theta}_2\right) &= Var\left(\widehat{\theta}_2\right) \\
&= Var\left(X - Y\right) \\
&= Var(X) + Var(Y) \\
&= 4 + 2 \\
&= 6.
\end{aligned}
$$

(ii) Both $\widehat{\theta}_1$ and $\widehat{\theta}_2$ are unbiased. The MSE of $\widehat{\theta}_1$ is smaller than the MSE of $\widehat{\theta}_2$. So, we prefer $\widehat{\theta}_1$.

(iii) The bias of $\widehat{\theta}_c$ is

$$
\begin{aligned}
E\left(\widehat{\theta}_c\right) - \theta &= E\left(\frac{c}{2}X + (1-c)Y\right) - \theta \\
&= \frac{c}{2}E(X) + (1-c)E(Y) - \theta \\
&= \frac{c}{2}2\theta + (1-c)\theta - \theta \\
&= 0,
\end{aligned}
$$

so $\widehat{\theta}_c$ is unbiased.

The variance of $\widehat{\theta}_c$ is

$$
\begin{aligned}
Var\left(\widehat{\theta}_c\right) &= Var\left(\frac{c}{2}X + (1-c)Y\right) \\
&= \frac{c^2}{4}Var(X) + (1-c)^2 Var(Y) \\
&= \frac{c^2}{4}4 + 2(1-c)^2 \\
&= c^2 + 2(1-c)^2.
\end{aligned}
$$

Let $g(c) = c^2 + 2(1-c)^2$. Them $g^{'}(c) = 6c - 4 = 0$ if $c = 2/3$. Also $g^{''}(c) = 6 > 0$. So, $c = 2/3$ minimizes the variance of $\widehat{\theta}_c$.

Solutions to Question 3 An electrical circuit consists of three batteries connected in series to a lightbulb. We model the battery lifetimes X_1, X_2, X_3 as independent and identically distributed $Exp(\lambda)$ random variables. Our experiment to measure the lifetime of the lightbulb Y is stopped when any one of the batteries fails. Hence, the only random variable we observe is $Y = \min(X_1, X_2, X_3)$.

(i) The cdf of Y is

$$\begin{aligned}
\Pr(Y < y) &= 1 - \Pr(Y > y) \\
&= 1 - \Pr\left[\min\left(X_1, X_2, X_3\right) > y\right] \\
&= 1 - \Pr\left(X_1 > y\right)\Pr\left(X_2 > y\right)\Pr\left(X_3 > y\right) \\
&= 1 - \Pr^3\left(X > y\right) \\
&= 1 - \exp(-3\lambda y).
\end{aligned}$$

So, $Y \sim Exp(3\lambda)$.

(ii) The likelihood function of λ is

$$L(\lambda) = 3\lambda \exp(-3\lambda y)$$

for $\lambda > 0$.

(iii) The log likelihood function of λ is

$$\log L(\lambda) = \log(3\lambda) - 3\lambda y.$$

The first derivative of $\log L$ with respect to λ is

$$\frac{d \log L(\lambda)}{d\lambda} = \frac{1}{\lambda} - 3y.$$

Setting this to zero and solving, we obtain $\widehat{\lambda} = 1/(3y)$. The second derivative of $\log L$ with respect to λ

$$\frac{d^2 \log L(\lambda)}{d\lambda^2} = -\frac{1}{\lambda^2} < 0,$$

so $\widehat{\lambda} = 1/(3y)$ is indeed a maximum likelihood estimator of λ.

(iv) The bias of $\widehat{\lambda}$ is

$$\begin{aligned}
Bias\left(\widehat{\lambda}\right) &= E\left(\widehat{\lambda}\right) - \lambda \\
&= \lambda \int_0^\infty \frac{1}{y} \exp(-3\lambda y) dy - \lambda \\
&= \infty,
\end{aligned}$$

so the estimator is biased.

(v) The mean squared error of $\widehat{\lambda}$ is

$$MSE\left(\widehat{\lambda}\right) = Var\left(\widehat{\lambda}\right) + Bias^2\left(\widehat{\lambda}\right) = Var\left(\widehat{\lambda}\right) + \infty = \infty.$$

Solutions to Question 4 Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with the common probability density function (pdf):

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}$$

for $x > 0$, $-\infty < \mu < \infty$ and $\sigma > 0$. Both μ and σ^2 are unknown.

(i) The joint likelihood function of μ and σ^2 is

$$\begin{aligned} L\left(\mu, \sigma^2\right) &= \prod_{i=1}^{n} \left\{ \frac{1}{X_i \sqrt{2\pi}\sigma} \exp\left[-\frac{(\log X_i - \mu)^2}{2\sigma^2}\right] \right\} \\ &= \frac{1}{(2\pi)^{n/2}\sigma^n} \left(\prod_{i=1}^{n} X_i^{-1}\right) \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (\log X_i - \mu)^2\right]. \end{aligned}$$

The joint log likelihood function of μ and σ^2 is

$$\log L\left(\mu, \sigma^2\right) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \sum_{i=1}^{n} \log X_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} (\log X_i - \mu)^2 .$$

The first order partial derivatives of this with respect to μ and σ are

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n} (\log X_i - \mu) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} \log X_i - n\mu\right) \tag{1}$$

and

$$\frac{\partial \log L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n} (\log X_i - \mu)^2 , \tag{2}$$

respectively.

(ii) Using equation (1), one can see that the solution of $\partial \log L/\partial\mu = 0$ is $\mu = \overline{\log X} = (1/n)\sum_{i=1}^{n} \log X_i$.

(iii) Using equation (2), one can see that the solution of $\partial \log L/\partial\sigma = 0$ is $\sigma^2 = (1/n)\sum_{i=1}^{n} (\log X_i - \widehat{\mu})^2$.

(iv) The mle, $\widehat{\mu}$, is an unbiased and consistent estimator for μ since

$$\begin{aligned} E\left(\widehat{\mu}\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n} \log X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E\left(\log X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mu \\ &= \mu \end{aligned}$$

and

$$\begin{aligned}
Var\left(\widehat{\mu}\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n}\log X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var\left(\log X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

(v) The mle, $\widehat{\sigma^2}$, is a biased and consistent estimator for σ^2 since

$$\begin{aligned}
E\left(\widehat{\sigma^2}\right) &= E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\log X_i - \widehat{\mu}\right)^2\right] \\
&= E\left[\frac{n-1}{n}S^2\right] \\
&= \frac{\sigma^2}{n}E\left[\frac{n-1}{\sigma^2}S^2\right] \\
&= \frac{\sigma^2}{n}E\left[\chi^2_{n-1}\right] \\
&= \frac{(n-1)\sigma^2}{n}
\end{aligned}$$

and

$$\begin{aligned}
Var\left(\widehat{\sigma^2}\right) &= Var\left[\frac{1}{n}\sum_{i=1}^{n}\left(\log X_i - \widehat{\mu}\right)^2\right] \\
&= Var\left[\frac{n-1}{n}S^2\right] \\
&= \frac{\sigma^4}{n^2}E\left[\frac{n-1}{\sigma^2}S^2\right] \\
&= \frac{\sigma^4}{n^2}E\left[\chi^2_{n-1}\right] \\
&= \frac{2(n-1)\sigma^4}{n^2}.
\end{aligned}$$

Note that we have used the fact $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Furthermore, $S^2 = (1/(n-1))\sum_{i=1}^{n}(X_i - \bar{X})^2$ denotes the sample variance.

Solutions to Question 5 Suppose we wish to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

(i) the Type I error occurs if H_0 is rejected when in fact $\theta = \theta_0$.

(ii) the Type II error occurs if H_0 is accepted when in fact $\theta \neq \theta_0$.

(iii) the significance level is the probability of type I error.

(iv) the power function: $\Pi(\theta) = \Pr(\text{ Reject } H_0 \mid \theta)$.

UP TO THIS BOOK WORK.

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\theta, \sigma^2)$, where σ^2 is assumed known.

(i) The power function, $\Pi(\theta)$, for $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is

$$
\begin{aligned}
\Pi(\theta) &= \Pr\left(\frac{\sqrt{n}}{\sigma} \mid \bar{x} - \theta_0 \mid > z_{\alpha/2} \middle| \theta\right) \\
&= \Pr\left(\mid \bar{x} - \theta_0 \mid > \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \middle| \theta\right) \\
&= \Pr\left(\bar{x} > \theta_0 + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \text{ or } \bar{x} < \theta_0 - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \middle| \theta\right) \\
&= \Pr\left(\bar{x} - \theta > \theta_0 - \theta + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \text{ or } \bar{x} - \theta < \theta_0 - \theta - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \middle| \theta\right) \\
&= \Pr\left(\frac{\sqrt{n}}{\sigma} (\bar{x} - \theta) > \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) + z_{\alpha/2} \text{ or } \frac{\sqrt{n}}{\sigma} (\bar{x} - \theta) < \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) - z_{\alpha/2} \middle| \theta\right) \\
&= \Pr\left(Z > \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) + z_{\alpha/2} \text{ or } Z < \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) - z_{\alpha/2}\right) \\
&= 1 - \Pr\left(Z < \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) + z_{\alpha/2}\right) + \Pr\left(Z < \frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) - z_{\alpha/2}\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) + z_{\alpha/2}\right) + \Phi\left(\frac{\sqrt{n}}{\sigma} (\theta_0 - \theta) - z_{\alpha/2}\right).
\end{aligned}
$$

(ii) The power function, $\Pi(\theta)$, for $H_0 : \theta = \theta_0$ versus $H_1 : \theta < \theta_0$ is

$$
\begin{aligned}
\Pi(\theta) &= \Pr\left(\frac{\sqrt{n}}{\sigma} (\bar{x} - \theta_0) \leq -z_\alpha \middle| \theta\right) \\
&= \Pr\left(\bar{x} - \theta_0 \leq -\frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right) \\
&= \Pr\left(\bar{x} \leq \theta_0 - \frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \Pr\left(\bar{x} - \theta \leq \theta_0 - \theta - \frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right) \\
&= \Pr\left(\frac{\sqrt{n}}{\sigma}(\bar{x} - \theta) \leq \frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) - z_\alpha \middle| \theta\right) \\
&= \Pr\left(Z \leq \frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) - z_\alpha\right) \\
&= \Phi\left(\frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) - z_\alpha\right).
\end{aligned}
$$

(iii) The power function, $\Pi(\theta)$, for $H_0 : \theta = \theta_0$ versus $H_1 : \theta > \theta_0$ is

$$
\begin{aligned}
\Pi(\theta) &= \Pr\left(\frac{\sqrt{n}}{\sigma}(\bar{x} - \theta_0) \geq z_\alpha \middle| \theta\right) \\
&= \Pr\left(\bar{x} - \theta_0 \geq \frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right) \\
&= \Pr\left(\bar{x} \geq \theta_0 + \frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right) \\
&= \Pr\left(\bar{x} - \theta \geq \theta_0 - \theta + \frac{\sigma}{\sqrt{n}} z_\alpha \middle| \theta\right) \\
&= \Pr\left(\frac{\sqrt{n}}{\sigma}(\bar{x} - \theta) \geq \frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) + z_\alpha \middle| \theta\right) \\
&= \Pr\left(Z \geq \frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) + z_\alpha\right) \\
&= 1 - \Pr\left(Z < \frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) + z_\alpha\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}}{\sigma}(\theta_0 - \theta) + z_\alpha\right).
\end{aligned}
$$

Note that we have used the fact $(\sqrt{n}/\sigma)(\bar{x} - \theta) = Z \sim N(0, 1)$. Furthermore, $\Phi(\cdot)$ denotes the cumulative distribution function of $N(0, 1)$.

Solutions to Question 6 The Neyman-Pearson test rejects $H_0 : \theta = \theta_1$ in favor of $H_1 : \theta = \theta_2$ if

$$\frac{L(\theta_1)}{L(\theta_2)} = \frac{\prod_{i=1}^{n} f(X_i; \theta_1)}{\prod_{i=1}^{n} f(X_i; \theta_2)} < k$$

for some k. UP TO THIS BOOK WORK.

Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform $(0, \theta)$ distribution.

(i) The most powerful test is to reject $H_0 : \theta = \theta_1$ if

$$\begin{aligned}
\frac{L(\theta_1)}{L(\theta_2)} &= \frac{\theta_1^{-n} I\{0 < X_1 < \theta_1\} I\{0 < X_2 < \theta_1\} \cdots I\{0 < X_n < \theta_1\}}{\theta_2^{-n} I\{0 < X_1 < \theta_2\} I\{0 < X_2 < \theta_2\} \cdots I\{0 < X_n < \theta_2\}} \\
&= \frac{\theta_2^n \, I\{\max(X_1, X_2, \ldots, X_n) < \theta_1\}}{\theta_1^n \, I\{\max(X_1, X_2, \ldots, X_n) < \theta_2\}} \\
&< k_0,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
& \frac{I\{\max(X_1, X_2, \ldots, X_n) < \theta_1\}}{I\{\max(X_1, X_2, \ldots, X_n) < \theta_2\}} < \frac{k_0 \theta_1^n}{\theta_2^n} \\
\iff & \max(X_1, X_2, \ldots, X_n) > k
\end{aligned}$$

as required. The last step follows because

$$\frac{I\{\max(X_1, X_2, \ldots, X_n) < \theta_1\}}{I\{\max(X_1, X_2, \ldots, X_n) < \theta_2\}}$$

is a decreasing function of $\max(X_1, X_2, \ldots, X_n)$.

(ii) The power function is

$$\begin{aligned}
\Pi(\theta) &= \Pr(\text{Reject } H_0 \mid \theta) \\
&= \Pr(\max(X_1, X_2, \ldots, X_n) > k \mid \theta) \\
&= 1 - \Pr(\max(X_1, X_2, \ldots, X_n) \leq k \mid \theta) \\
&= 1 - \Pr(X_1 \leq k \mid \theta) \Pr(X_2 \leq k \mid \theta) \cdots \Pr(X_n \leq k \mid \theta) \\
&= 1 - \left(\frac{k}{\theta}\right)^n.
\end{aligned}$$

(iii) Note that

$$\begin{aligned}
& 1 - (2k)^5 = 0.05 \\
\iff & 2k = (0.95)^{1/5} \\
\iff & k = (1/2)(0.95)^{1/5}.
\end{aligned}$$

So, $k = 0.4948969$.

(iv) Note that

$$\begin{aligned}
\beta &= \Pr(\text{Type II error}) \\
&= 1 - \left[1 - \left(\frac{0.4948969}{0.6}\right)^5\right] \\
&= \left(\frac{0.4948969}{0.6}\right)^5 \\
&= 0.3817837.
\end{aligned}$$