Problem 1.

Answer: a)

$$\alpha(c) = P(T \in C; H_0) = P(X > c; \mu = 0)$$
$$= P\left(\frac{X - 0}{1} > \frac{c - 0}{1}\right)$$
$$= P(Z > c) = 1 - \Phi(c)$$

b)

$$\beta(c) = P(T \notin C; \mu = 1)$$
$$= P\left(\frac{X - 1}{1} \leq \frac{c - 1}{1}\right)$$
$$= P(Z \leq c - 1) = \Phi(c - 1)$$

c) Given

$$R(c) = 1 - \Phi(c) + \Phi(c - 1)$$
$$R'(c) = -\Phi'(c) + \Phi'(c - 1)$$
$$= -\frac{\exp\{-c^2/2\}}{\sqrt{2\pi}} + \frac{\exp\{-(c-1)^2/2\}}{\sqrt{2\pi}}$$

Let $R'(c) = 0$,

$$-\frac{\exp\{-c^2/2\}}{\sqrt{2\pi}} + \frac{\exp\{-(c-1)^2/2\}}{\sqrt{2\pi}} = 0$$
$$\Rightarrow \exp\{-c^2/2\} = \exp\{-(c-1)^2/2\}$$
$$\Rightarrow c^2 = (c - 1)^2$$
$$\Rightarrow c^2 = c^2 - 2c + 1 \Rightarrow 2c = 1, c = \frac{1}{2}$$

Then

$$R''(c) = -\frac{c \exp\{-c^2/2\}}{\sqrt{2\pi}} + \frac{(1 - c) \exp\{-(c-1)^2/2\}}{\sqrt{2\pi}}$$
$$R''(c = 1/2) = 0.352065 > 0 \text{ (local min) and,}$$
$$R(1/2) = 1 - \Phi(1/2) + \Phi(-1/2)$$
$$= 2\Phi(-1/2)$$

Hence, $c^* = \frac{1}{2}, R(c^*) = 2\Phi(-1/2)$ □

Problem 2.

Answer: a) Given the likelihood ratio:

$$\frac{L(1)}{L(2)} \le k$$

$$\Rightarrow \frac{(2\pi)^{-n/2} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i^2)\right]}{(2\pi \cdot 2)^{-n/2} \exp\left[-\frac{1}{2\cdot 2}\sum_{i=1}^{n}(x_i^2)\right]}$$

$$= 2^{n/2} \exp\left[-\frac{1}{4}\sum_{i=1}^{n}(x_i^2)\right]$$

$$\Rightarrow \ln 2^{n/2} - \frac{1}{4}\sum_{i=1}^{n}(x_i^2) \le \ln k$$

$$\Rightarrow \frac{1}{4}\sum_{i=1}^{n}(x_i^2) \ge \frac{1}{2}\ln 2^n - \frac{1}{2}\ln k^2$$

$$\Rightarrow \sum_{i=1}^{n} x_i^2 \ge 2\ln \frac{2^{10}}{k^2} = c$$

By Neyman-Pearson Lemma, the best critical region is of the form:

$$C = \left\{ (x_1, \ldots x_{10}); \sum_{i=1}^{n} x_i^2 \ge c \right\}$$

where c is determined as

$$0.05 = P\left(\sum_{i=1}^{n} x_i^2 \ge c; \sigma^2 = 1\right)$$

$$\Rightarrow 0.05 = P\left(\chi^2(10) \ge c; \sigma^2 = 1\right)$$

$$= 1 - P\left(\chi^2(10) \le c\right)$$

$$\Rightarrow P\left(\chi^2(10) \le c\right) = 0.95$$

$$\Rightarrow c = \chi^2_{0.05}(10) = 18.31$$

Hence for n = 10, the test with critical region $\{\sum_{i=1}^{n} x_i^2 \ge 18.31\}$ is the most powerful test of size $\alpha = 0.05$.
b) From R, $K(2) = 1 - P\left(\frac{1}{2}\sum_{i=1}^{n} x_i < \frac{18.41}{2}\right) = 0.5174595$

c) Given the log-likelihood ratio:

$$\frac{L(1)}{L(4)} \le k$$

$$\Rightarrow \frac{(2\pi)^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i^2)\right]}{(2\pi \cdot 4)^{-n/2}\exp\left[-\frac{1}{2\cdot 4}\sum_{i=1}^{n}(x_i^2)\right]}$$

$$= 2^n \exp\left[-\frac{3}{8}\sum_{i=1}^{n}(x_i^2)\right]$$

$$\Rightarrow \ln 2^n - \frac{3}{8}\sum_{i=1}^{n}x_i^2 \le \ln k$$

$$\Rightarrow \sum_{i=1}^{n}x_i^2 \ge \frac{8}{3}\ln\frac{2^{10}}{k} = c$$

By Neyman-Pearson Lemma, the best critical region is of the form:

$$C = \left\{(x_1,\ldots x_{10}); \sum_{i=1}^{n}x_i^2 \ge c\right\}$$

where c is determined as

$$\Rightarrow c = \chi^2_{0.05}(10) = 18.31$$

Hence for n = 10, the test with critical region $\{\sum_{i=1}^{n}x_i^2 \ge 18.31\}$
is the most powerful test of size $\alpha = 0.05$.
d) From R, $K(4) = 1 - P\left(\frac{1}{4}\sum_{i=1}^{n}x_i < \frac{18.41}{4}\right) = 0.917559$
e) Given the likelihood ratio:

$$\frac{L(1)}{L(\sigma_1^2)} \le k$$

$$\Rightarrow \frac{(2\pi)^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i^2)\right]}{(2\pi \cdot \sigma_1^2)^{-n/2}\exp\left[-\frac{1}{2\cdot\sigma_1^2}\sum_{i=1}^{n}(x_i^2)\right]}$$

$$= \sigma_1^n \exp\left[-\frac{\sigma_1^2-1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i^2)\right]$$

$$\Rightarrow \ln\sigma_1^n - \frac{\sigma_1^2-1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i^2) \le \ln k$$

$$\Rightarrow \sum_{i=1}^{n}x_i^2 \ge \frac{2\sigma_1^2}{\sigma_1^2-1}\ln\frac{\sigma_1^{10}}{k} = c$$

By Neyman-Pearson Lemma, the best critical region is of the form:

$$C = \left\{(x_1,\ldots x_{10}); \sum_{i=1}^{n}x_i^2 \ge c\right\}$$

3

where c is determined as

$$\Rightarrow c = \chi_{0.05}^2(10) = 18.31$$

Hence for n = 10, the test with critical region $\{\sum_{i=1}^n x_i^2 \geq 18.31\}$
is the most powerful test of size $\alpha = 0.05$.

f) Since the critical region C defines a test that is most powerful against each simple alternative $\sigma^2 > 1$, the test in (e) is a uniformly most powerful test, and C is a uniformly most powerful critical region of size $\alpha = 0.05$. Again, if $\alpha = 0.05$, then $c = \chi_{0.05}^2(10) = 18.31$. $\square$

Problem 3.

Answer: a)

$$\bar{X} \sim N\left(\mu_1, \frac{400}{n}\right), \bar{Y} \sim N\left(\mu_2, \frac{225}{n}\right)$$

$$\Rightarrow \bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{400}{n} + \frac{225}{n} = \frac{625}{n}\right)$$

$$\Rightarrow \bar{X} - \bar{Y} \sim N\left(\theta, \frac{625}{n}\right)$$

The power function is given as:

$$K(\theta) = P(\bar{x} - \bar{y} \in C; \theta)$$

$$= P\left(\frac{\bar{x} - \bar{y} - \theta}{\sqrt{625/n}} \geq \frac{c - \theta}{\sqrt{625/n}}\right)$$

$$= 1 - \Phi\left(\frac{c - \theta}{25\sqrt{n}}\right)$$

b) Given $\alpha = 0.05, \theta = 10$, we have

$$0.05 = \alpha(H_0 : \theta = 0) = P(\bar{x} - \bar{y} \in C; \theta = 0)$$

$$\Rightarrow 0.05 = 1 - \Phi\left(\frac{c}{25/\sqrt{n}}\right)$$

$$\Rightarrow \left(\frac{c}{25/\sqrt{n}}\right) = 1.645 \text{ and,}$$

$$0.9 = K(10) = P(\bar{x} - \bar{y} \in C; \theta = 10)$$

$$\Rightarrow 0.9 = 1 - \Phi\left(\frac{c - 10}{25/\sqrt{n}}\right)$$

$$\Rightarrow \left(\frac{c - 10}{25/\sqrt{n}}\right) = -1.28$$

Solving for n and c: $c = 5.62393, n = 54.4782 \approx 54$ $\square$

Problem 4.

Answer: a) The likelihood function for X and Y, respectively, are:

$$L_X(\mu_1) = \frac{1}{(\sqrt{2\pi})^{m/2}} \exp\left\{-\frac{\sum_{i=1}^m (x_i - \mu_1)^2}{2}\right\},$$

$$L_Y(\mu_2) = \frac{1}{(\sqrt{2\pi})^{n/2}} \exp\left\{-\frac{\sum_{j=1}^n (y_j - \mu_2)^2}{2}\right\}$$

Then

$$L(\mu_1, \mu_2) = L_X(\mu_1)L_Y(\mu_2)$$

$$= \frac{1}{(2\pi)^{m/2}} \exp\left\{-\frac{\sum_{i=1}^m (x_i - \mu_1)^2}{2}\right\} \cdot \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{\sum_{j=1}^n (y_j - \mu_2)^2}{2}\right\}$$

$$= \frac{1}{(2\pi)^{(m+n)/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^m (x_i - \mu_1)^2\right\} \exp\left\{-\frac{1}{2}\sum_{j=1}^n (y_j - \mu_2)^2\right\}$$

b) We find the gradient of of the log-likelihood function $l(\mu_1, \mu_2)$ and set it to $\mathbf{0}$

$$\left(\partial l(\mu_1,\mu_2)/\partial \mu_1 \quad \partial l(\mu_1, \mu_2)/\partial \mu_2\right)^T$$

$$= \begin{pmatrix} \partial\left[\ln \frac{\exp\left\{-\frac{1}{2}\sum_{j=1}^n (y_j-\mu_2)^2\right\}}{(2\pi)^{(m+n)/2}} - \frac{1}{2}\sum_{i=1}^m (x_i - \mu_1)^2\right]/\partial\mu_1 \\ \partial\left[\ln \frac{\exp\left\{-\frac{1}{2}\sum_{i=1}^m (x_i-\mu_1)^2\right\}}{(2\pi)^{(m+n)/2}} - \frac{1}{2}\sum_{j=1}^n (y_j - \mu_2)^2\right]/\partial\mu_2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \sum_{i=1}^m (x_i - \mu_1) \\ \sum_{j=1}^n (y_j - \mu_2) \end{pmatrix} = \mathbf{0}$$

$$\Rightarrow \begin{pmatrix} \sum_{i=1}^m x_i = m\mu_1 \\ \sum_{j=1}^n y_j = n\mu_2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \hat{\mu}_1 = \frac{1}{m}\sum_{i=1}^m x_i \\ \hat{\mu}_2 = \frac{1}{n}\sum_{j=1}^n y_j \end{pmatrix}$$

The Hessian of the $l(\mu_1, \mu_2)$ can be found as:

$$H_l(\mu_1, \mu_2) = \begin{pmatrix} \partial^2 l(\mu_1,\mu_2)/\partial \mu_1^2 & \partial^2 l(\mu_1,\mu_2)/\partial \mu_1 \partial \mu_2 \\ \partial^2 l(\mu_1,\mu_2)/\partial \mu_2 \partial \mu_1 & \partial^2 l(\mu_1,\mu_2)/\partial \mu_2^2 \end{pmatrix} \Rightarrow \begin{pmatrix} -m & 0 \\ 0 & -n \end{pmatrix}$$

$$\Rightarrow \det(\lambda I - H_l(\mu_1, \mu_2)) = 0$$

$$\Rightarrow \det\begin{pmatrix} \lambda + m & 0 \\ 0 & \lambda + n \end{pmatrix} \Rightarrow (\lambda + m)(\lambda + n) = 0$$

$$\lambda = -m, \lambda = -n, \text{ for all } n, m > 0$$

The Hessian and of $l(\mu_1, \mu_2)$ has negative eigenvalues $\lambda_1 = -m < 0$ and $\lambda_2 = -n < 0$ and is thus negative definite. Hence $\left(\hat{\mu}_1 = \bar{X} \quad \hat{\mu}_2 = \bar{Y}\right)^T$ is a local and global maximizer of

5

the log-likelihood function, and thus $\hat{\mu}_{1,mle} = \frac{1}{m}\sum_{i=1}^{m}X_i = \bar{X}$, $\hat{\mu}_{2,mle} = \frac{1}{n}\sum_{j=1}^{n}Y_i = \bar{Y}$.

c) Since under $H_0 : \mu_1 = \mu_2$, we find

$$\frac{dl(\mu)}{d\mu} = \frac{d\left[\ln\frac{1}{(\sqrt{2\pi})^{(m+n)/2}} - \frac{1}{2}\sum_{i=1}^{m}(x_i - \mu)^2 - \frac{1}{2}\sum_{j=1}^{n}(y_j - \mu)^2\right]}{d\mu}$$

$$\Rightarrow \sum_{i=1}^{m}(x_i - \mu) + \sum_{j=1}^{n}(y_j - \mu) = 0$$

$$\Rightarrow m\bar{x} - m\mu + n\bar{y} - n\mu = 0$$

$$\Rightarrow \mu(m + n) = m\bar{x} + n\bar{y}$$

$$\Rightarrow \mu = \frac{m\bar{x} + n\bar{y}}{m + n}$$

Hence $\hat{\mu}_{mle} = \frac{m\bar{X}+n\bar{Y}}{m+n}$

d) We find

$$max_{\mu=\mu_1=\mu_2}L(\mu) = \frac{1}{(\sqrt{2\pi})^{(m+n)/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \hat{\mu}_{mle})^2\right\}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(y_j - \hat{\mu}_{mle})^2\right\}$$

$$max_{\mu_1,\mu_2}L(\mu_1, \mu_2) = \frac{1}{(\sqrt{2\pi})^{(m+n)/2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \bar{x})^2\right\}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(y_j - \bar{y})^2\right\}$$

$$\Rightarrow \frac{max_{\mu=\mu_1=\mu_2}L(\mu)}{max_{\mu_1,\mu_2}L(\mu_1, \mu_2)} = \frac{\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \hat{\mu}_{mle})^2 - \frac{1}{2}\sum_{j=1}^{n}(y_j - \hat{\mu}_{mle})^2\right\}}{\exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \bar{x})^2 - \frac{1}{2}\sum_{j=1}^{n}(y_j - \bar{y})^2\right\}}$$

$$= \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \hat{\mu}_{mle})^2 - \frac{1}{2}\sum_{j=1}^{n}(y_j - \hat{\mu}_{mle})^2 + \frac{1}{2}\sum_{i=1}^{m}(x_i - \bar{x})^2 + \frac{1}{2}\sum_{j=1}^{n}(y_j - \bar{y})^2\right\}$$

Since

$$\sum_{i=1}^{m}(x_i - \hat{\mu})^2 = \sum_{i=1}^{m}(x_i^2 - 2x_i\hat{\mu} + \hat{\mu}^2)$$

$$= \sum_{i=1}^{m}(x_i^2) - 2m\bar{x}\hat{\mu} + m\hat{\mu}^2,$$

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

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

$$\Rightarrow -\sum_{i=1}^{m}(x_i - \hat{\mu})^2 + \sum_{i=1}^{m}(x_i - \bar{x})^2$$

$$= 2m\bar{x}\hat{\mu} - m\hat{\mu}^2 - m\bar{x}^2$$

Similarly

$$-\sum_{j=1}^{n}(y_j - \hat{\mu})^2 + \sum_{j=1}^{n}(y_j - \bar{y})^2$$

$$= 2n\bar{y}\hat{\mu} - n\hat{\mu}^2 - n\bar{x}^2$$

Then

$$-\sum_{i=1}^{m}(x_i - \hat{\mu})^2 + \sum_{i=1}^{m}(x_i - \bar{x})^2 - \sum_{j=1}^{n}(y_j - \hat{\mu})^2 + \sum_{j=1}^{n}(y_j - \bar{y})^2$$

$$= 2m\bar{x}\hat{\mu} - m\hat{\mu}^2 - m\bar{x}^2 + 2n\bar{y}\hat{\mu} - n\hat{\mu}^2 - n\bar{x}^2$$

$$= (2m\bar{x} + 2n\bar{y})\frac{m\bar{x} + n\bar{y}}{m+n} - (m+n)\left(\frac{m\bar{x} + n\bar{y}}{m+n}\right)^2 - (m\bar{x}^2 + n\bar{y}^2)\left(\frac{m+n}{m+n}\right)$$

$$= \frac{mn}{m+n}(2\bar{x}\bar{y} - x^2 - y^2 = -\frac{mn}{m+n}(\bar{x} - \bar{y})^2$$

Hence

$$\frac{max_{\mu=\mu_1=\mu_2}L(\mu)}{max_{\mu_1,\mu_2}L(\mu_1,\mu_2)}$$

$$= \exp\left\{-\frac{1}{2}\sum_{i=1}^{m}(x_i - \hat{\mu}_{mle})^2 - \frac{1}{2}\sum_{j=1}^{n}(y_j - \hat{\mu}_{mle})^2 + \frac{1}{2}\sum_{i=1}^{m}(x_i - \bar{x})^2 + \frac{1}{2}\sum_{j=1}^{n}(y_j - \bar{y})^2\right\}$$

$$= \exp\left\{-\frac{1}{2}\frac{mn}{(m+n)}(\bar{x} - \bar{y})^2\right\}$$

e) Under $H_0 : \mu_1 = \mu_2$, the distribution of $\bar{X} - \bar{Y}$ and test statistic is given as:

$$\bar{X} - \bar{Y} \sim N\left(0, \frac{1}{m} + \frac{1}{n}\right)$$

$$\Rightarrow T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

$$\Rightarrow T^2 = \frac{(\bar{X} - \bar{Y})^2}{\frac{m+n}{mn}}$$

$$= \frac{mn}{m+n}(\bar{X} - \bar{Y})^2$$

Set $\lambda \le k$

$$\exp\left\{-\frac{1}{2}\frac{mn}{(m+n)}(\bar{x} - \bar{y})^2\right\} \le k$$

$$\Rightarrow -\frac{t^2}{2} \le \ln k$$

$$\Rightarrow t^2 \ge 2\ln k$$

$$\Rightarrow |t| \ge \sqrt{2\ln k} = c$$

Hence the critical region is of the form

$$C = \left\{ |t| = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{1}{m} + \frac{1}{n}}} \geq c \right\}$$

To determine c, let

$$\alpha = P(|T| \geq c; \mu_1 = \mu_2) = P(|Z| \geq c)$$

$$\Rightarrow P\left(\left| \frac{\bar{x} - \bar{y}}{\sqrt{\frac{1}{m} + \frac{1}{n}}} \right| \geq z_{\alpha/2} \right) = \alpha$$

$$\Rightarrow C = \left\{ \sqrt{\frac{mn}{m+n}} \, |\bar{x} - \bar{y}| \geq z_{\alpha/2} \right\}$$

$\square$