Carson Slater

Dr. Harvill

STA 5353

Homework 5

1) Given $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 known. For $H_0: \mu \leq \mu_0$ and $H_1: \mu > \mu_0$, the test rejects for all $\bar{x} > \mu_0 + K$.

a) The power function is:

$$\beta(\mu) = P(\bar{x} > \mu_0 + K) = 1 - P(\bar{x} < \mu_0 + K)$$

$$= 1 - P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{(\mu_0 + K) - \mu}{\sigma/\sqrt{n}}\right)$$

$$= 1 - \phi\left(\frac{(\mu_0 + K) - \mu}{\sigma/\sqrt{n}}\right)$$

$$= \phi\left(-\frac{(\mu_0 + K) - \mu}{\sigma/\sqrt{n}}\right). \square$$

b) See plot at end of homework.

c) A valid p-value is such that

$$p(\underline{x}) = \sup_{\mu \leq \mu_0} P(\bar{X} \geq \bar{x}).$$

We reject H_0 if $\bar{X} > \mu + K$ for some K. So

$$p(\underline{x}) = P_{\mu_0}(\bar{X} < \bar{x}) = \phi\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right)$$

is a valid p-value.

2) Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, σ^2 known.

a) We know $\ell(\mu|\underline{x}, \sigma^2) = \log(L(\mu|\underline{x}, \sigma^2))$ is

$$\ell(\mu|x, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

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

Maximizing w.r.t. μ, we have

$$0 \overset{set}{=} \frac{\partial}{\partial\mu}\ell(\mu|x, \sigma^2) = -\frac{1}{2\sigma^2}\left(2n\mu + 2n\bar{x}\right)$$

$$\Rightarrow \hat{\mu}_{MLE} = \bar{x}$$

To confirm this is a global maximum,

$$\frac{\partial^2}{\partial\mu^2}\ell(\mu|x, \sigma^2) = \frac{\partial}{\partial\mu} -\frac{1}{2\sigma^2}\left(2n\mu + 2n\bar{x}\right) = -n < 0.$$

So $\hat{\mu} = \bar{x}$, is the MLE for μ, as $\ell(\mu|\underline{x}, \sigma^2)$ is concave down $\forall\mu$.

b) Consider $H_0: \mu = \mu_0$ versus $H_1: \mu \neq \mu_0$

We need to find the LRT statistic.

$$\lambda(\underline{x}) = \frac{L(\mu_0|\underline{x})}{L(\bar{x}|\underline{x})} = \frac{(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2\right)\right\}}{(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)\right\}}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x})^2\right)\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left(n(\bar{x}-\mu_0)^2\right)\right\} \quad \left(b/c \ \sum_{i=1}^{n}(x_i-\mu_0)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2\right)$$

The LRT is the test that rejects the null hypothesis for $\left\{\underline{x} : \exp\left\{-\frac{1}{2\sigma^2}\left(n(\bar{x}-\mu_0)^2\right)\right\} \le c\right\}$. See plot at end of homework.

c) The LRT rejection region can be simplified:

$$\left\{\underline{x} : \log\left(\exp\left\{-\frac{n}{2\sigma^2}(\bar{x}-\mu_0)^2\right\}\right) \le \log(c)\right\}$$

$$= \left\{\underline{x} : (\bar{x}-\mu_0)^2 \ge -\frac{2\sigma^2}{n}\log(c)\right\}$$

$$= \left\{\underline{x} : |\bar{x}-\mu_0| \ge \sqrt{-\frac{2\sigma^2}{n}\log(c)}\right\}$$

$$= \left\{\underline{x} : |\bar{x}-\mu_0| \ge K\right\}, \text{ where } K = \sqrt{-\frac{2\sigma^2}{n}\log(c)}.$$

So the LRT is the test that rejects H_0 when $\bar{x}$ and μ_0 differ by more than a specified amount.

d) $\beta(\theta) = P(|\bar{x}-\mu_0| \ge K) = P(\bar{x}-\mu_0 < -K \cup \bar{x}-\mu_0 > K)$

$$= P(\bar{x} < \mu_0 - K \cup \bar{x} > \mu_0 + K)$$

$$= P\left(\bar{x} < \mu_0 - K\right) + \left(1 - P\left(\bar{x} < \mu_0 + K\right)\right)$$

$$= P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{(\mu_0 - K) - \mu}{\sigma/\sqrt{n}}\right) + 1 - P\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} < \frac{(\mu_0 + K) - \mu}{\sigma/\sqrt{n}}\right)$$

$$= \phi\left(\frac{(\mu_0 - K) - \mu}{\sigma/\sqrt{n}}\right) + 1 - \phi\left(\frac{(\mu_0 + K) - \mu}{\sigma/\sqrt{n}}\right) \leftarrow \text{power fn.}$$

See plot at end of homework.

3) Given a random sample $Y_1, \ldots, Y_m \overset{iid}{\sim} \text{Bernoulli}(\theta)$, we consider to test $H_0: \theta = 0.49$ verses $H_1: \theta = 0.51$

a) Find the UMP level α test.

Using the Neyman-Pearson Lemma, the test that will reject Ho for all $\underline{y}$ such that

$$f(\underline{y} \mid \theta = 0.51) > k f(\underline{y} \mid \theta = 0.49)$$

$$\frac{f(\underline{y} \mid \theta = 0.51)}{f(\underline{y} \mid \theta = 0.49)} > k$$

$$\frac{(0.51)^{\sum_{i=1}^m Y_i} (0.49)^{m - \sum_{i=1}^m Y_i}}{(0.49)^{\sum_{i=1}^m Y_i} (0.51)^{m - \sum_{i=1}^m Y_i}} > k \qquad \left(\frac{51}{49}\right)^{w} \left(\frac{51}{49}\right)^{w-m}$$

$$\left(\frac{49}{51}\right)^{m - 2\sum_{i=1}^m Y_i} > k \qquad \left(\frac{51}{49}\right)^{2w-m}$$

$$\left(\frac{49}{51}\right)^{m-2w} > k$$

The UMP level α test is the test that rejects H_0 if $\left(\frac{49}{51}\right)^{m-2w} > k$.

$$\Rightarrow w \leq c$$

b) See table at end of homework.

c) If $W \sim \text{Bin}(m, \theta)$, then by the CLT, with sufficiently large sample size, $W \sim N(m\theta, m\theta(1-\theta))$

To find the error probabilities,

$$P\left(\left(\frac{49}{51}\right)^{m-2w} > k\right) = P(W \leq c) \quad \text{where } c = \tfrac{1}{2}\left(\log\left(\tfrac{49}{51}\right) - \log k + m\right)$$

Probs are equal

We want to solve $P_{H_0}(W \leq c) = 0.01$ and $P_{H_1}(W \leq c) = 0.01$

So then by CLT $\quad$ qnorm(0.01, 0, 1)

$$\frac{c - m(0.49)}{\sqrt{m(0.49)(0.51)}} = -2.33 \quad \text{and} \quad \frac{c - m(0.51)}{\sqrt{m(0.51)(0.49)}} = 2.33$$

$$c = m(0.49) - 2.33\sqrt{m(0.49)(0.51)}$$

$$\Rightarrow \frac{\left(m(0.49) - 2.33\sqrt{m(0.49)(0.51)}\right) - m(0.51)}{\sqrt{m(0.49)(0.51)}} = 2.33$$

$$\Rightarrow m = 13{,}567$$

$$\Rightarrow c = 6512.161$$

So the sample size m must be $m \geq 13567$ and decision rule $W \leq 6512$.

d) Consider $\quad H_0: \theta \leq \frac{1}{2}$ versus $H_1: \theta > \frac{1}{2}$

To apply the Karlin-Rubin Theorem, $W = \sum_{i=1}^{m} Y_i$ must be sufficient for θ. By the Factorization Theorem, we have that

$$f(\underline{y} \mid \theta) = \theta^{\sum_{i=1}^{m} y_i}(1-\theta)^{\sum_{i=1}^{m} y_i}, \quad y_i = 0, 1.$$

$$= g(T(\underline{y}) \mid \theta)\, h(\underline{y})$$

Where $g(T(\underline{y}) \mid \theta) = f(\underline{y} \mid \theta)$ and $h(\underline{y}) = 1$. So $T(\underline{y})$ is sufficient for θ.

We must also show the Binomial family has monotone likelihood ratio. For $\theta_2 > \theta_1$,

$$\frac{g(w|\theta_2)}{g(w|\theta_1)} = \frac{\binom{n}{w}\theta_2^w(1-\theta_2)^{n-w}}{\binom{n}{w}\theta_1^w(1-\theta_1)^{n-w}}$$

$$= \left(\frac{\theta_2(1-\theta_1)}{\theta_1(1-\theta_2)}\right)^w \left(\frac{1-\theta_2}{1-\theta_1}\right)^n.$$

Since $0<\theta_i<\theta_2<1$, $\frac{g(w|\theta_2)}{g(w|\theta_1)}$ is increasing $\forall$ possible values of $\theta_1 < \theta_2$, since $\frac{\theta_2}{\theta_1} > 1$ and $\frac{1-\theta_1}{1-\theta_2} > 1$. So the binomial family has MLR. By the Karlin-Rubin Theorem, the UMP test is to reject H_0 if $W > t_0$ where t_0 satisfies $\alpha \geq P_{\theta_0}(W \geq w_0)$.

e)

w_0	0	1	2	3	4	5	6	7	8	9
size	0.99	0.98	0.94	0.82	0.62	0.37	0.17	0.054	0.010	0.0009
								0.1	0.05	0.01

Corresponds to α level

f) A valid p-value is a p-value $p(\underline{x})$ such that

$$p(\underline{x}) = \sup_{\theta\in\Theta_0} P_\theta\left(W(\underline{x}) \geq W(\underline{x})\right)$$

$$= \sum_{W(\underline{x})}^{m} \binom{m}{w}\left(\tfrac{1}{2}\right)^w\left(\tfrac{1}{2}\right)^{m-w}$$

$$= \sum_{W(\underline{x})}^{m} \binom{m}{w}\left(\tfrac{1}{2}\right)^m$$

is a valid p-value. If $w=7$, the p-value is 0.05468.

4) a) Given a normal family, σ^2 known. Also $\mu_2 > \mu_1$.

$$\frac{g(x|\mu_2)}{g(x|\mu_1)} = \frac{\cancel{\frac{1}{\sqrt{2\pi\sigma^2}}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu_2)^2\right\}}{\cancel{\frac{1}{\sqrt{2\pi\sigma^2}}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu_1)^2\right\}}$$

$$= \frac{\exp\left\{-\cancel{\frac{x^2}{2\sigma^2}} + \frac{x\mu_2}{\sigma^2} - \frac{\mu_2^2}{\sigma^2}\right\}}{\exp\left\{-\cancel{\frac{x^2}{2\sigma^2}} + \frac{x\mu_1}{\sigma^2} - \frac{\mu_1^2}{\sigma^2}\right\}}$$

$$= \exp\left\{\frac{x(\mu_2-\mu_1)}{\sigma^2} + \frac{\mu_1^2-\mu_2^2}{2\sigma^2}\right\}$$

Since the likelihood ratio is increasing in x $\forall$ $\mu_2 > \mu_1$, then the family $N(\mu, \sigma^2)$ has MLR.

b) Given a Poisson family with rate λ. Let $\lambda_2 > \lambda_1$

$$\frac{g(x|\lambda_2)}{g(x|\lambda_1)} = \frac{\frac{\lambda_2^x e^{-\lambda_2}}{x!}}{\frac{\lambda_1^x e^{-\lambda_1}}{x!}} = \left(\frac{\lambda_2}{\lambda_1}\right)^x e^{\lambda_1-\lambda_2}.$$

This is increasing in x $\forall$ $\lambda_2 > \lambda_1$, so this Poisson family has MLR.

5) Let $g(t|\theta) = h(t)c(\theta)\exp\{w(\theta)t\}$ be a one-parameter exponential family. Also, let $\theta_2 > \theta_1$.

The Likelihood ratio is

$$\frac{g(t|\theta_2)}{g(t|\theta_1)} = \frac{\cancel{h(t)}c(\theta_2)\exp\{w(\theta_2)t\}}{\cancel{h(t)}c(\theta_1)\exp\{w(\theta_1)t\}}$$

$$= \frac{c(\theta_2)}{c(\theta_1)}\exp\{(w(\theta_2)-w(\theta_1))t\}$$

This function is monotone given $w(\cdot)$ is monotone over values of θ. So then this exponential family has MLR.

Three examples of one-parameter exponential families are $\text{Exp}(\lambda)$, $\text{Gamma}(1, \beta)$, and $N(\mu, 1)$.

6) Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, 1)$, let θ_0 be a specified value of Θ. Consider $H_0: \theta \geq \theta_0$ vs. $H_1: \theta < \theta_0$.

a) In homework 1, problem 2, we showed that $\bar{X}$ is a sufficient statistic for θ. On page 113 of Casella + Berger, we have that a $N(\mu, \sigma^2)$ is is an exponential family, so $N(\theta, 1)$ is a single parameter exponential family. We know $\bar{X} \sim N(\theta, 1/\sqrt{n})$ where N is known, making the distribution of $T(\underline{x})$ a single parameter exponential family. In Problem 5 of this homework, we showed a single parameter exponential family has MLR. So the family of pdfs of $T(\underline{x})$ has MLR. So by the Karlin-Rubin theorem, for any t_0, the test that rejects H_0 if and only if $T < K$ is a UMP level α test such that $\alpha = P(\bar{X} < K)$.

b) We consider that an unbiased test is such that $\beta(\theta') \geq \beta(\theta'')$ for $\theta' \in \Theta_0^c$ and $\theta'' \in \Theta_0$. We have that

$$\beta(\theta) = P(\bar{X} \in R) = P(\bar{X} < K) = \Phi\left(\frac{K - \theta}{\sigma/\sqrt{n}}\right) \left(\text{since } \bar{X} \sim N(\theta, \tfrac{\sigma^2}{n})\right)$$

This is the standard normal CDF, and is mono-tonically increasing over its support. Considering the hypotheses $H_0: \theta'' \geq \theta_0$ and $H_1: \theta' < \theta_0$

this implies $\theta'' > \theta'$. So then $B(\theta'') \geq B(\theta')$.

as $B(\theta) = \phi(\cdot)$ is monotonically increasing.

So this test is not unbiased.

c) The MLE of θ: $\underset{\text{LLH}}{\frac{\partial}{\partial\theta} \ell(\theta|\underline{x})} = \frac{\partial}{\partial\theta}\left[-\frac{n}{2}(2\pi) - \frac{1}{2}\sum_{i=1}^{n} x_i^2 + \theta\sum_{i=1}^{n} x_i - \frac{n\theta^2}{2}\right]$

$$0 \overset{\text{set}}{=} \sum_{i=1}^{n} x_i - n\theta \implies \hat{\theta}_{MLE} = \bar{x}$$

Check Global Max:

$$\frac{\partial^2}{\partial\theta} \ell(\theta|x)\Big|_{\theta=\hat{\theta}} = \frac{\partial}{\partial\theta}\sum_{i=1}^{n} x_i - n\theta = -n.$$

$\ell(\theta)$ Concave down $\forall\,\theta$ so $\hat{\theta} = \bar{x}$ is global max.

The LRT statistic is

$$\lambda(\underline{x}) = \frac{L(\theta_0|\underline{x})}{L(\hat{\theta}|\underline{x})} = \begin{cases} \exp\left\{n\bar{x}(\theta_0 - \hat{\theta}) + \frac{\theta_0^2 - \hat{\theta}^2}{2}\right\} & \theta_0 < \hat{\theta} \\ 1 & \theta_0 \geq \hat{\theta}. \end{cases}$$

The LRT is the test that rejects H_0 when

$\{\underline{x} : \lambda(\underline{x}) < c\} \ni 0 \leq c \leq 1$. Also then when $\theta_0 < \hat{\theta}$,

this is equivalent to $\left\{\underline{x} : \bar{x} < \frac{\left(\log(c) - \frac{\theta_0^2 - \hat{\theta}^2}{2}\right)}{n(\theta_0 - \hat{\theta})}\right\}$

which is the same as $\bar{x} < K$ where

$$K = \frac{\left(\log(c) - \frac{\theta_0^2 - \hat{\theta}^2}{2}\right)}{n(\theta_0 - \hat{\theta})}.$$

same test as (a)

d) For any specified α value, a level α test satisfies

$$P(X \in R) \leq \alpha .$$

Consider an alternative point $\theta_1 < \theta_0$. From our result in (b) we can show that the test rejects H_0 if $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$, which is $\sup_{\theta \in \Theta_0} P(X \in R)$. By the necessity principle of the Neyman-Peason Lemma, any other α level test that has as high a power at θ_1 for $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ must have the same rejection region. So if UMP level α test exists, it must be this test $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ as no other test has a higher power at θ_1.

Now consider $\theta_2 > \theta_0$ which rejects H_0 if $\bar{X} > \theta_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$. Then the power of this test is,

$$\beta(\theta_2) = P_{\theta_2}\left(\bar{X} > \theta_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right) = P\left(\frac{\bar{X}-\theta_2}{\sigma/\sqrt{n}} > z_\alpha + \frac{\theta_0 - \theta_2}{\sigma/\sqrt{n}}\right)$$

since $\theta_0 - \theta_2 = 0$

$$> P(Z > z_\alpha) = P(Z < -z_\alpha)$$

$$> P_{\theta_2}\left(\frac{\bar{X}-\theta_2}{\sigma/\sqrt{n}} < -z_\alpha + \frac{\theta_0 - \theta_2}{\sigma/\sqrt{n}}\right) = P\left(\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

Where $P\left(\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}\right)$ is the power of $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$.

So then since the power of $\bar{X} > \theta_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$ is greater than the power of $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ at $\theta_2 > 0$, and $\bar{X} < \theta_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ has the highest power at $\theta_1 < 0$, then an UMP level α test does not exists.

7) Consider $H_0: p \le 1/3$ and $H_1: p > 1/3$.

Let $X \sim \text{Binomial}(5, p)$. For 0-1 loss, $c_I = c_{II} = 1$ for

$$L(\theta, a_0) = \begin{cases} 0 & p \in P_0 \\ c_I & p \in P_0^c \end{cases} \text{ and } L(\theta, a_0) = \begin{cases} c_I & p \in P_0 \\ 0 & p \in P_0^c. \end{cases}$$

$$R(\theta, \delta(\underline{X})) = \begin{cases} \beta(p) & 0 \le p \le 1/3 \\ 1 - \beta(p) & 1/3 < p < 1 \end{cases}$$

For test 1: $\beta(p) = P(X \le 1)$

$$= \sum_{i=0}^{1} \binom{5}{x} (p)^x (1-p)^{5-x}$$

For test 2: $\beta(p) = P(X \ge 4)$

$$= \sum_{i=4}^{5} \binom{5}{x} (p^x)(1-p)^{5-x}$$

The Risk for test 1 is higher for most extreme values of p, and the risk for test 2 is only higher for values $p \in (1/3, 2/3)$.

So test 2 is preferred.

See plot at end of homework

8) Let $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, σ^2 known, τ^2 known, $\mu = 0$.

Consider $H_0: \theta \le 0$ vs. $H_1: \theta > 0$.

a) We know $P(\theta \le 0 \mid \underset{\sim}{x})$ is the probability H_0 is true. So then

$$P(\theta < 0 \mid \underset{\sim}{x}) = P\left(Z \le \frac{0 - E[\theta \mid \underset{\sim}{x}]}{\sqrt{Var(\theta \mid \underset{\sim}{x})}} \,\middle|\, \bar{x}\right)$$

$$= P\left(Z \le \frac{-\frac{\tau^2 \bar{x} +}{\tau^2 + \sigma^2/n}}{\sqrt{\frac{\sigma^2 \tau^2/n}{\sigma^2/n + \tau^2}}}\right)$$

$$= P\left(Z \le \frac{-\sqrt{n}\,\tau \bar{x}}{\sigma \sqrt{\tau^2 + \sigma^2/n}}\right)$$

$$= \phi\left(\frac{-\sqrt{n}\,\tau \bar{x}}{\sigma \sqrt{\tau^2 + \sigma^2/n}}\right).$$

b) By the definition of a p-value,

$$p(\underset{\sim}{x}) = \sup_{\theta \in \Theta_0} P(\bar{X} < \bar{x} \mid \theta = 0) = P\left(Z < \frac{0 - \sqrt{n}\,\bar{x}}{\sigma}\right)$$

Sub $\mu = \bar{x}$

$$= \phi\left(-\frac{\sqrt{n}\,\bar{x}}{\sigma}\right) \text{ since } \bar{X} \sim N(\mu, \sigma^2/n)$$

c) Let $\sigma^2 = \tau^2 = 1$. So then

$$P(\theta < 0) = \phi\left(\frac{-\sqrt{n}\,\tau\,\bar{x}}{\sigma\sqrt{\tau^2 + \sigma^2/n}}\right) \quad \text{from (a).}$$

$$= \phi\left(\frac{-\sqrt{n}\,\bar{x}}{1\sqrt{1 + 1/n}}\right) \quad \text{by substitution.}$$

$$= \phi\left(\frac{-n\bar{x}}{\sqrt{n+1}}\right).$$

$$P(\underline{x}) = \phi\left(-\frac{\sqrt{n}\,\bar{x}}{\sigma}\right) = \phi\left(-\sqrt{n}\,\bar{x}\right).$$

We know the standard normal CDF ϕ is monotone increasing over its domain. We have that $\frac{-n\bar{x}}{\sqrt{n+1}} > -\sqrt{n}\,\bar{x}$. So the the probability the posterior is true is always greater than the p-value.

d) Here we have

$$\lim_{\tau^2 \to \infty} P(\theta \le 0 \mid \underline{x}) = \lim_{\tau^2 \to \infty} \phi\left(\frac{-\sqrt{n}\,\tau\,\bar{x}}{\sigma\sqrt{\tau^2 + \sigma^2/n}}\right)$$

$$= \lim_{\tau^2 \to \infty} \phi\left(\frac{-\sqrt{n}\;\bar{x}}{\sigma\sqrt{1 + \sigma^2/\tau^2 n}}\right) = \phi\left(\frac{-\sqrt{n}\,\bar{x}}{\sigma\sqrt{1+0}}\right) = \phi\left(\frac{-\sqrt{n}\,\bar{x}}{\sigma}\right) = p(\underline{x}).$$

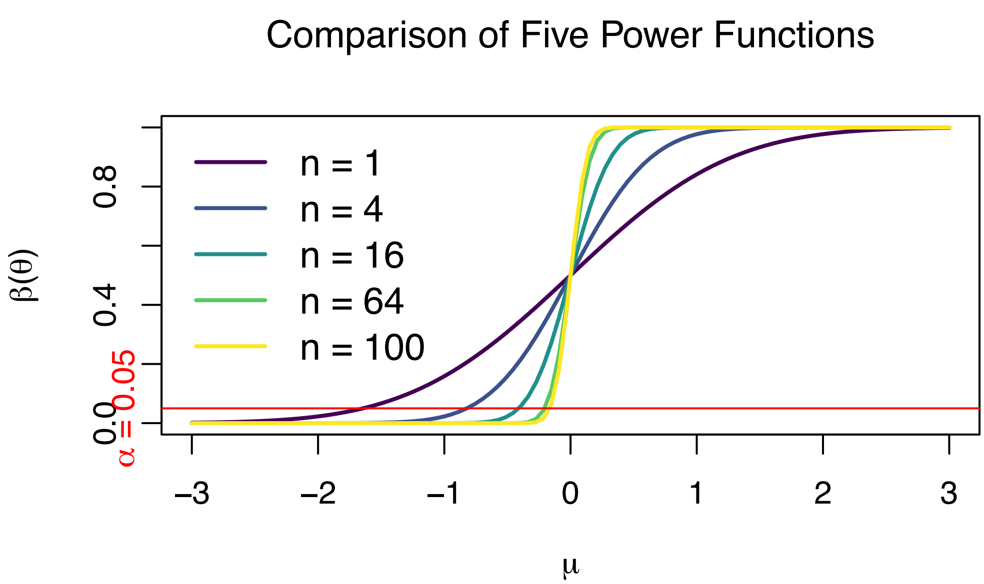
Theory 2 Homework 5 R Output

Carson Slater

1(b)

```
n <- c(1, 4, 16, 64, 100)
mu <- seq(-3, 3, length = 100)
mu_0 <- 0
sigma <- 1
K <- 0
beta <- list()

# colors
cols <- viridis::viridis(5)

# power function varying by sample size
for (i in 1:5) beta[[i]] <- pnorm(-(mu_0 + K - mu)/((sigma/sqrt(n[i]))))

# plot
plot(mu, beta[[1]], type = "l", lwd = 2, col = cols[1],
     main = latex2exp::TeX(r'(Comparison of Five Power Functions)'),
     xlab = latex2exp::TeX(r'(\mu)'),
     ylab = latex2exp::TeX(r'(\beta(\theta))'))
for (i in 2:5) lines(mu, beta[[i]], col = cols[i], lwd = 2)
abline(h = 0.05, col = "red")
mtext(latex2exp::TeX(r'(\alpha = 0.05)'), side = 2,
      line = 0.5, col = "red", at = c(0.05))
legend("topleft", c("n = 1",
                     "n = 4",
                     "n = 16",
                     "n = 64",
                     "n = 100"),
       col = cols,
       lwd = c(2, 2, 2, 2, 2),
```

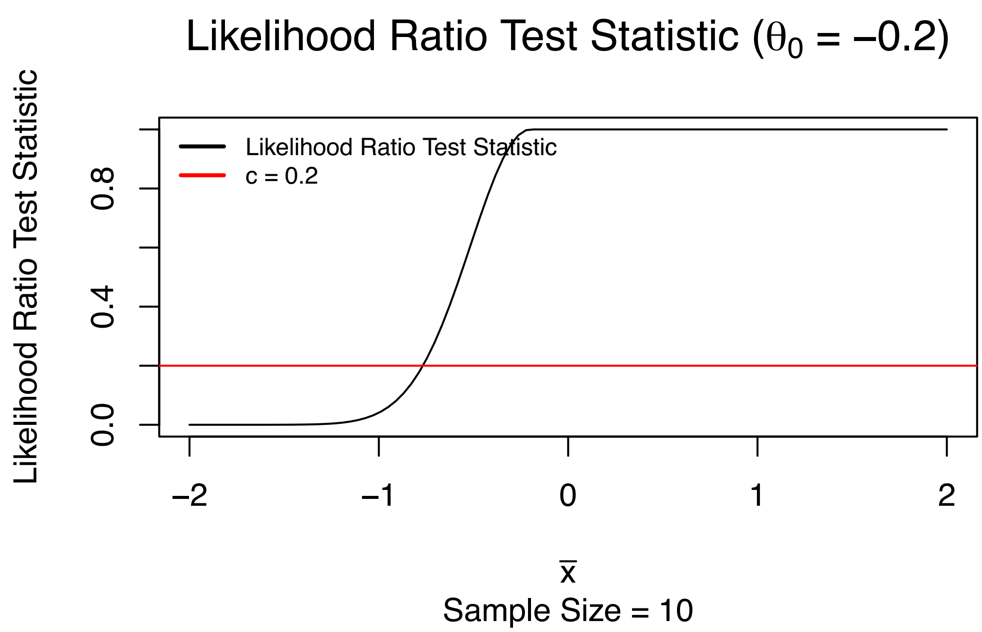
```
       bty = "n",
       cex = 1.2)
```



2(b)

```
xbar <- seq(-2, 2, length.out = 100)
sigma_sq <- 1
theta_0 <- -0.2
n <- 10

lrt <- function(xbar, theta_0, sigma_sq, n) {
  output <- exp(-n*(xbar - theta_0)^2/(2*sigma_sq))
  for (i in seq_along(xbar)) {
    if (xbar[i] >= theta_0) output[i] <- 1
  }
  output
}

lrt <- lrt(xbar, theta_0 = -0.2, sigma_sq = 1, n = 10)

lam <- exp(-n*(xbar - theta_0)^2/(2*sigma_sq))

plot(xbar, lrt, type = "l", col = "black",
```

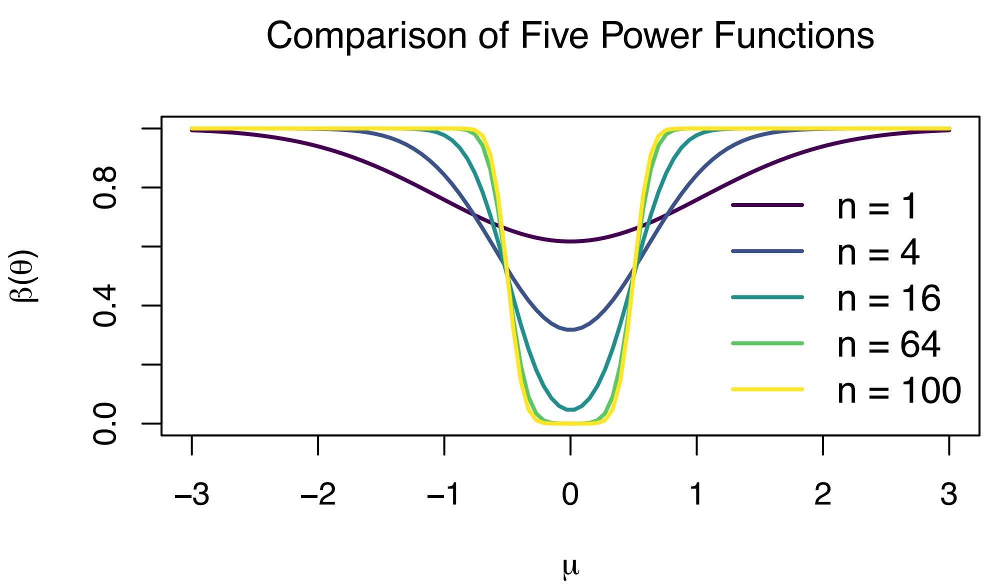
```
     xlab = latex2exp::TeX(r'(\bar{x})'),
     ylab = "Likelihood Ratio Test Statistic",
     xlim = c(-2, 2),
     main = latex2exp::TeX(r'(Likelihood Ratio Test Statistic ($\theta_0$ = -0.2))'),
     cex.main = 1.3)
title(sub = "Sample Size = 10")
abline(h = 0.2, col = "red")
legend("topleft", c("Likelihood Ratio Test Statistic",
                    "c = 0.2"),
       col = c("black", "red"),
       lwd = c(2, 2),
       lty = c(1, 1),
       bty = "n",
       cex = 0.75)
```



2(d)

```
n <- c(1, 4, 16, 64, 100)
mu <- seq(-3, 3, length = 100)
mu_0 <- 0
sigma <- 1
K <- 0.5 # K is now 0.5
```

```
beta <- list()

# colors
cols <- viridis::viridis(5)

# power function varying by sample size
for (i in 1:5) beta[[i]] <- pnorm((mu_0 - K - mu)/((sigma/sqrt(n[i])))) + 1 - pnorm((mu_0 + K

# plot
plot(mu, beta[[1]], type = "l", lwd = 2, col = cols[1],
     main = latex2exp::TeX(r'(Comparison of Five Power Functions)'),
     xlab = latex2exp::TeX(r'(\mu)'),
     ylab = latex2exp::TeX(r'(\beta(\theta))'),
     ylim = c(0,1))
for (i in 2:5) lines(mu, beta[[i]], col = cols[i], lwd = 2)
legend("bottomright", c("n = 1",
                        "n = 4",
                        "n = 16",
                        "n = 64",
                        "n = 100"),
       col = cols,
       lwd = c(2, 2, 2, 2, 2),
       bty = "n",
       cex = 1.2)
```



3(b)

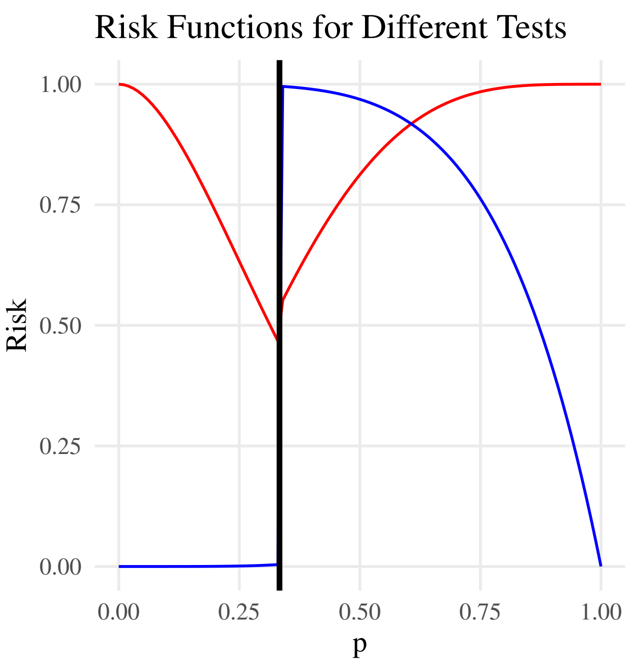
```r
m <- 10
w <- 1:10
k <- sapply(w, \(w) (51/49)^(2*w - m))
# since W is Binomial
# P(X in R | Theta = 0.49)
prob <- pbinom(w, size = m, 0.49, lower.tail = FALSE)

cols <- c("W", "k", "Probability Sum Y's is In R")

cbind(w, k, prob) |> as.data.frame() |> knitr::kable(col.names = cols)
```

W	k	Probability Sum Y's is In R
1	0.7261180	0.9873722
2	0.7866027	0.9379222
3	0.8521256	0.8112268
4	0.9231065	0.5982047
5	1.0000000	0.3526028
6	1.0832986	0.1559607
7	1.1735359	0.0480003
8	1.2712898	0.0091028
9	1.3771865	0.0007979
10	1.4919043	0.0000000

7

```r
R1 <- function(p) ifelse(p <= 1/3, pbinom(1, 5, p), 1 - pbinom(1, 5, p))
R2 <- function(p) ifelse(p <= 1/3, pbinom(4, 5, p, lower.tail = FALSE), 1 - pbinom(4, 5, p, 

p <- seq(0, 1, length.out = 101)
r1 <- R1(p)
r2 <- R2(p)

ggplot() +
  geom_line(aes(p, r1), color = "red") +
  geom_line(aes(p, r2), color = "blue") +
  geom_vline(xintercept = 1/3,
```

```
                 color = "black", size=1) +
  labs(title = "Risk Functions for Different Tests",
       y = "Risk") +
  coord_equal()
```



The risk for test 1 is the red line. The risk for test 2 is the blue line. The black line was added at $p = \frac{1}{3}$.