STAT 135 HW 3 Vinay Maruri

1] (1) $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ ← definition

$Bias(\hat{\theta}_n) = E(\hat{\theta}_n - \theta)$ ← definition
$= E(\hat{\theta}_n) - \theta$

$Bias(\bar{X}_n) = E(\bar{X}_n - \lambda) = E(\bar{X}_n) - \lambda$

$E(\bar{X}_n) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \lambda$

$= \frac{1}{n} \cdot n\lambda = \lambda$ expectation of poisson r.v. is λ.

Hence, $Bias(\bar{X}_n) = E(\bar{X}_n) - \lambda = \lambda - \lambda = 0$

$E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right) \cancel{= \frac{1}{n-1}\sum_{i=1}^{n} E[(X_i - \bar{X}_n)^2]}$

$\cancel{= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2 + \bar{X}_n^2 - 2X_i\bar{X}_n)}$

$\cancel{= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) + \sum_{i=1}^{n} E(\bar{X}_n^2)\right.}$

$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 + n\bar{X}_n^2 - 2\bar{X}_n \sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}_n^2)\right]$

$= \frac{1}{n-1}\left[n(\mu^2 + \sigma^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right]$ ← using second moments

$= \frac{1}{n-1}(n\sigma^2 - \sigma^2) = \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2$

variance of poisson r.v. is λ.

$\Rightarrow E(s^2) = \sigma^2 = \lambda$

Hence, $Bias(s^2) = E(s^2 - \lambda) = E(s^2) - \lambda = \lambda - \lambda = 0$.

Thus, sample mean and sample variance here are unbiased estimators of λ.

② From lecture 6: (slides 3+4)

candidate unbiased estimators for θ: $\hat{\theta}_n, \breve{\theta}_n, \mathring{\theta}_n, \ldots$

$E_\theta(\hat{\theta}_n) = \theta$ and CR bound $= \frac{1}{nI(\theta)}$

For $X_1, \ldots, X_n$ iid poisson (λ)

$I(\lambda) = \frac{1}{\lambda}$, CR bound $= \frac{1}{nI(\lambda)} = \frac{1}{n} \cdot \frac{1}{\frac{1}{\lambda}}$

CR bound $= \frac{\lambda}{n}$

③ by theorem E of lecture 6 [slide 3]:

cramer-rao lower bound, $\rightarrow Var(s^2) \geq \frac{\lambda}{n}$ $Var(\bar{X}_n) \geq \frac{\lambda}{n}$

since s^2 + $\bar{X}_n$ are unbiased estimators of λ.

by definition: $Var(\bar{X}_n) = \frac{\sigma^2}{n}$ where σ^2 is variance of a poisson r.v

$Var(\bar{X}_n) = \frac{\lambda}{n}$

since $\bar{X}_n$ achieves cramer-rao lower bound, it is a UMVUE of λ.

From question, $Var(s^2) = \frac{E(X-\mu)^4}{n} - \frac{\sigma^4(n-3)}{n(n-1)}$

and $E(X-\lambda)^4 = \lambda(1+3\lambda)$

$\sigma^2 = \lambda$

$$Var(s^2) = \frac{\lambda(1+3\lambda)}{n} - \frac{\sigma^4(n-3)}{n(n-1)} = \frac{\lambda(1+3\lambda)}{n} - \frac{\lambda^2(n-3)}{n(n-1)}$$

$$= \frac{\lambda + 3\lambda^2}{n} - \frac{\lambda^2(n-3)}{n(n-1)} = \frac{(\lambda+3\lambda^2)(n-1) - \lambda^2 n + 3\lambda^2}{n(n-1)}$$

$$= \frac{n\lambda - \lambda + 3n\lambda^2 - \cancel{3\lambda^2} - \lambda^2 n \cancel{+3\lambda^2}}{n(n-1)}$$

$$= \frac{\lambda(n-1) + n\lambda^2(3-1)}{n(n-1)} = \frac{\lambda(n-1) + 2n\lambda^2}{n(n-1)}$$

$$= \frac{\lambda(n-1)}{n(n-1)} + \frac{2n\lambda^2}{n(n-1)} = \frac{\lambda}{n} + \frac{2\lambda^2}{n-1}$$

$$Var(s^2) = \frac{\lambda}{n} + \frac{2\lambda^2}{n-1}$$

Hence s^2 doesn't meet the cramer-rao lower bound and is not a UMVUE of λ.

2] $X \sim f(x|\theta_0)$. show: $E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(X|\theta_0)\right] = 0$.

① $E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(X|\theta_0)\right] = E_{\theta_0}\left[\frac{\frac{\partial}{\partial\theta} f(x|\theta_0)}{f(x|\theta_0)}\right]$ use exchangeability

$= \int_{\theta_0} \frac{\frac{\partial}{\partial\theta} f(x|\theta_0)}{f(x|\theta_0)} f(x|\theta_0)\,dx = \int_{\theta_0} \frac{\partial}{\partial\theta} f(x|\theta_0)\,dx$

$= \frac{\partial}{\partial\theta}\int_{\theta_0} f(x|\theta_0)\,dx = \frac{\partial}{\partial\theta}[1] = 0$

② $f(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$ is joint density

$E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(X_1,\ldots,X_n|\theta_0)\right] = E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log \prod_{i=1}^{n} f(x_i|\theta_0)\right]$

$= E_{\theta_0}\left[\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x_i|\theta_0)\right] = E_{\theta_0}\left[\sum_{i=1}^{n}\frac{\frac{\partial}{\partial\theta} f(x_i|\theta_0)}{f(x_i|\theta_0)}\right]$

$= \sum_{i=1}^{n}\int_{\theta_0}\frac{\frac{\partial}{\partial\theta} f(x_i|\theta_0)}{f(x_i|\theta_0)} f(x_i|\theta_0)\,dx_i$

use exchangeability

$= \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\int_{\theta_0} f(x_i|\theta_0)\,dx_i$

$= \sum_{i=1}^{n}\frac{\partial}{\partial\theta}[1] = \sum_{i=1}^{n} 0 = 0$

Hence,
$E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(x_1,\ldots,x_n|\theta_0)\right] = 0$

③ show: $E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(X_1,\ldots,X_n|\theta_0)\right]^2 = n I(\theta_0)$

$E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(x_1,\ldots,x_n|\theta_0)\right]^2 = E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log\prod_{i=1}^{n} f(x_i|\theta_0)\right]^2$

$= E_{\theta_0}\left[\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(x_i|\theta_0)\right]^2 = \sum_{i=1}^{n} E_{\theta_0}\left[\frac{\partial}{\partial\theta}\log f(x_i|\theta_0)\right]^2$

definition: $I(\theta) = E_\theta\left[\frac{\partial}{\partial\theta}\log f(X|\theta)\right]^2$

thus $\Rightarrow \sum_{i=1}^{n} I(\theta_0) = n I(\theta_0)$

3] ① $L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^n$

$\cancel{\ell(\theta) = \log L(\theta) = \log\left(\left(\frac{1}{\theta}\right)^n\right) = n(\log 1 - \log\theta) = n\log 1 - n\log\theta}$

$\cancel{\frac{\partial \ell(\theta)}{\partial\theta} = -\frac{n}{\theta}}$

$0 \le x_i \le \theta, \; i=1,\dots,n$

$\iff 0 \le X_{(1)} < X_{(n)} \le \theta$

$\iff \theta \ge X_{(n)}, \; \theta \ge X_{(1)}, \; \theta \ge 0$

$\iff \theta \ge \max\{0, X_{(1)}, X_{(n)}\}$

with this $L(\theta)$, likelihood is maximized with biggest possible θ, hence $\hat{\theta}_{MLE} = \max\{0, X_{(1)}, X_{(n)}\}$

This is a sufficient statistic for θ:

- lecture 6, slide 7 says using $X_{(1)}$ and $X_{(n)}$ to infer θ is sufficient
- By factorization theorem, this is true: Let $X_n = (x_1, \dots, x_n)$

$$f(X_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \left(\frac{1}{\theta}\right)^n \mathbb{1}\{\theta \ge \max\{0, X_{(1)}, X_{(n)}\}\}$$

where $\mathbb{1}\{\theta \ge \max(0, X_{(1)}, X_{(n)})\} = \begin{cases} 1 & \text{if true} \\ 0 & \text{if false} \end{cases}$

hence, $f(X_n|\theta) = g[\hat{\theta}_{MLE}, \theta] \cdot h(X_n)$

where $h(X_n) = 1$

$$g[\hat{\theta}_{MLE}, \theta] = \left(\frac{1}{\theta}\right)^n \mathbb{1}\{\theta \ge \max\{0, X_{(1)}, X_{(n)}\}\}$$

hence $\hat{\theta}_{MLE}$ is a sufficient statistic for θ.

② $f(x|\theta) = \begin{cases} 1/\theta & \text{if } x \in (0,\theta) \\ 0 & \text{otherwise} \end{cases}$

$\log f(x|\theta) = \log\left(\frac{1}{\theta}\right) = \log(1) - \log(\theta) = -\log(\theta)$

$\frac{\partial \log f}{\partial \theta} = -\frac{1}{\theta}$ $\quad \frac{\partial^2 \log f}{\partial \theta^2} = \frac{1}{\theta^2}$

Hence, $I(\theta) = -\frac{1}{\theta^2}$

③ No we can't apply Theorem D of lectures because of the exchangeability assumptions for differentiation/integration.

$\frac{\partial}{\partial \theta} \log f(x|\theta) = -\frac{1}{\theta}$ $\quad \int \frac{\partial}{\partial \theta} \log f(x|\theta)\, dx$

$= \int_0^\theta -\frac{1}{\theta}\, dx = -\frac{1}{\theta}\int_0^\theta dx$

$\frac{\partial}{\partial \theta} \int \log f(x|\theta)\, dx = \frac{\partial}{\partial \theta}\int_0^\theta -\log(\theta)\, dx$ $\quad = -\frac{1}{\theta}[\theta]$

$= \frac{\partial}{\partial \theta}\left[-\log(\theta)\int_0^\theta dx\right]$ $\quad = -1$

$= \frac{\partial}{\partial \theta}\left[-\theta \log(\theta)\right] = -\theta\left(\frac{1}{\theta}\right) - \log(\theta)$

$= -1 - \log(\theta).$

Hence, $\int \frac{\partial}{\partial \theta} \log f(x|\theta)\, dx \neq \frac{\partial}{\partial \theta} \int \log f(x|\theta)\, dx$

④ From (1), $\hat{\theta}_{MLE} = \max\{0, X_{(1)}, X_{(n)}\}$.

Let $X_{(n)}$ be the sample maximum; then,

$\hat{\theta}_{MLE} = X_{(n)}$

Let $X_{(n)} = \max(X_1, \ldots, X_n)$ where X_i are as specified in question.

$$P(X_{(n)} \le x) = P(X_1 \le x, \ldots, X_n \le x) = [F_X(x)]^n$$

From problem, the density of $X_{(n)}$ is given by:

$$f_{X_{(n)}}(x) = n f(x) F^{n-1}(x)$$

Since $X_{(n)}$ is drawn from a uniform $(0,\theta)$ distribution [inferred from question] its CDF is given by: $\leftarrow$ CDF is $\frac{x}{\theta}$

$$F_{X(n)}(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \left(\frac{x}{\theta}\right)^n & x \in (0,\theta) \\ 1 & \text{if } x \ge \theta \end{cases}$$

Thus, $f_{X(n)}(x) = n\left[\frac{1}{\theta}\right]\left[\frac{x}{\theta}\right]^{n-1} = n \cdot \frac{x^{n-1}}{\theta^n}$

This is a ~~a~~ Beta $(n,1)$ random variable!

pdf: $\frac{x^{n-1}(1-x)^{1-1}}{\text{Beta}(n,1)} = \frac{x^{n-1}}{\frac{\Gamma(n)\Gamma(1)}{\Gamma(n+1)}} = x^{n-1}\left[\frac{\Gamma(n+1)}{\Gamma(n)\Gamma(1)}\right] = x^{n-1}\left[\frac{n!}{(n-1)!}\right] = n x^{n-1}$

from probability review in week 1, $U_{(n)} = \max(U_1, U_2, \ldots, U_n) \sim \text{Beta}(n,1)$

~~Hence, we need to find θ:~~

$\int f_{X(n)} x \, dx$ $\leftarrow$ shows that integral of pdf equals 1.

$$\int_0^\theta \frac{n \cdot x^{n-1}}{\theta^n} dx$$

[confirms intuition that this is a Beta $(n,1)$ random variable, but over $U(0,\theta)$ instead of $U(0,1)$]

$$= \frac{n}{\theta^n} \int_0^\theta x^{n-1} dx$$

$$= \frac{n}{\theta^n}\left[\frac{1}{n}x^n\right]_0^\theta = \frac{1}{\theta^n}[x^n]_0^\theta = \frac{1}{\theta^n}[\theta^n - 0]$$

$$= 1$$

Since the sampling distribution is Beta(n, 1), it can't be approximated by a normal distribution.

(5) $E(Y) = E\left(\frac{n+1}{n} X_{(n)}\right) = \frac{n+1}{n} E[X_{(n)}]$

$= \frac{n+1}{n} \int_0^\theta x \cdot n \frac{x^{n-1}}{\theta^n} dx = \frac{n+1}{n} \cdot \frac{n}{\theta^n} \int_0^\theta x \cdot x^{n-1} dx$

$= \frac{n+1}{\theta^n} \int_0^\theta x^n dx = \frac{n+1}{\theta^n} \left[\frac{1}{n+1} x^{n+1}\right]_0^\theta$

$= \frac{1}{\theta^n} \left[\theta^{n+1} - 0\right] = \frac{\theta^{n+1}}{\theta^n} = \theta$

Hence, $Bias(Y) = E(Y - \theta) = E(Y) - \theta = 0$
(Y is unbiased estimator.)

$Var(Y) = Var\left(\frac{n+1}{n} X_{(n)}\right) = \left(\frac{n+1}{n}\right)^2 Var(X_{(n)})$

$Var(X_{(n)}) = \int_0^\theta x^2 n \frac{x^{n-1}}{\theta^n} dx - [\theta]^2$

$= \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx - \theta^2$

$= \frac{n}{\theta^n} \left[\frac{1}{n+2} x^{n+2}\right]_0^\theta - \theta^2$

$= \left(\frac{n}{n+2} \cdot \frac{\theta^{n+2}}{\theta^n}\right) - \theta^2 = \frac{n}{n+2}(\theta^2) - \theta^2$

$Var(X_{(n)}) = \theta^2 \left(\frac{n}{n+2} - 1\right)$

$$Var(Y) = \left(\frac{n+1}{n}\right)^2 \left[\theta^2\left(\frac{n}{n+2} - 1\right)\right]$$

CR-bound: $\frac{1}{nI(\theta)} = \frac{1}{n(-\frac{1}{\theta^2})} = \frac{1}{-\frac{n}{\theta^2}} = -\frac{\theta^2}{n}$

$$Var(Y) = \left(\frac{n^2+2n+1}{n^2}\right)\left(\frac{\theta^2 n}{n+2} - \theta^2\right)$$

$$= \frac{\theta^2 n(n^2+2n+1)}{n^2(n+2)} - \frac{\theta^2(n^2+2n+1)}{n^2}$$

$$= \frac{\theta^2 n(n^2+2n+1) - \theta^2(n^2+2n+1)(n+2)}{n^2(n+2)}$$

$$= \frac{(n-(n+2))(\theta^2(n^2+2n+1))}{n^2(n+2)}$$

$$= \frac{-2(\theta^2(n^2+2n+1))}{n^2(n+2)} = -\frac{\theta^2}{n}\left[\frac{2(n^2+2n+1)}{n(n+2)}\right]$$

$$Var(Y) = -\frac{\theta^2}{n}\left[\frac{2n^2+4n+2}{n^2+2n}\right] \quad >1$$

since numerator larger than denominator

Hence $Var(Y)$ is smaller than the CR bound.

This result is not contradictory since the CR inequality assumptions don't hold in this case. [in part (3) of this question, I showed that the exchangeability assumption fails in this case. lecture 5, slide 20 states exchangeability as one of the assumptions for the cramer-rao inequality.}

4] (1) From samples in question, I wrote code to compute following:

$\bar{X}_n = 3.657$

$t_{13}(0.005) = 3.012276$

$S_x^2 = 39.480$

$\chi^2_{13}(0.005) = 29.81947$

$\hat{\sigma}_x^2 = 36.660$

$\chi^2_{13}(1-0.005) = 3.565035$

99% CI for μ:

$$\bar{X}_n \pm \frac{s}{\sqrt{n}} t_{n-1}\left(\frac{\alpha}{2}\right) = 3.657 \pm \left(\sqrt{\frac{39.480}{14}}\right) t_{13}\left(\frac{0.01}{2}\right)$$

$$= 3.657 \pm \left(\sqrt{2.82}\right)(3.012276)$$

$$= [-1.401, 8.715]$$

99% CI for σ^2:

$$\left[\frac{n\hat{\sigma}_n^2}{\chi^2_{n-1}\left(\frac{\alpha}{2}\right)} \le \sigma^2 \le \frac{n\hat{\sigma}_n^2}{\chi^2_{n-1}\left(1-\frac{\alpha}{2}\right)}\right]$$

$$= \left[\frac{14(36.660)}{29.81947}, \frac{14(36.660)}{3.565035}\right]$$

$$= [17.212, 143.965]$$

② $\bar{X}_n = 3.657,\ S_x^2 = 39.480,\ \hat{\sigma}_x^2 = 36.660$

99% CI (bootstrap) for μ:

$$\bar{X}_n \pm \frac{z_{\alpha/4}}{\sqrt{nI(\mu)}} \qquad I(\mu,\sigma^2) = \begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{bmatrix}$$

$$3.657 \pm \frac{z_{0.01/4}}{\sqrt{14 \cdot \frac{1}{\sigma^2}}}$$ ← qnorm(1−0.0025) in R.

$$3.657 \pm \frac{2.807}{\sqrt{14 \cdot \frac{1}{39.480}}} = 3.657 \pm 4.714$$

$$= [-1.057, 8.371]$$

99% CI (bootstrap) for σ^2:

$$\hat{\sigma}_n \pm \frac{z_{\alpha/4}}{\sqrt{nI(\sigma)}} = \sqrt{36.660} \pm \frac{2.807}{\sqrt{14 \cdot \frac{2}{36.660}}}$$

$$= 6.055 \pm \frac{2.807}{\sqrt{0.7638}}$$

$$= 6.055 \pm 3.2119$$

$$= [2.8431, 9.2669]^2$$

final CI for σ^2 $\Rightarrow [8.083, 85.875]$

③ The bootstrap confidence intervals have more "accuracy" because they have tighter confidence intervals than the exact confidence intervals.

④ I think we would need a sample twice as large to halve the length of the confidence interval for μ.

5] $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$; $H_0: p = 0.59$

$H_1: p = 0.45$

① From lecture 7, slide 22:

let R be rejection region of an LRT.

$$R = \{\lambda(X_n) \le c\}, \quad c \in [0,1]$$

$$\lambda(X_n) = \frac{\sup_{\Theta_0} L(\Theta | X_n)}{\sup_{\Theta} L(\Theta | X_n)}$$

$$L(p | \bar{X}_n) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$$

$$\lambda(\bar{X}_n) = \frac{\sup_{H_0} L(p | \bar{X}_n)}{\sup_{H} L(p | \bar{X}_n)} = \frac{L(0.59 | \bar{X}_n)}{\max\{L(0.59 | \bar{X}_n), L(0.45 | \bar{X}_n)\}}$$

$$= \frac{0.59^{\Sigma x_i} \, 0.41^{n-\Sigma x_i}}{\max\{0.59^{\Sigma x_i} \, 0.41^{n-\Sigma x_i}, \, 0.45^{\Sigma x_i} \, 0.55^{n-\Sigma x_i}\}}$$

$$= \frac{1}{\max\left\{1, \left(\frac{0.45}{0.59}\right)^{\Sigma x_i}\left(\frac{0.55}{0.41}\right)^{n-\Sigma x_i}\right\}}$$

$$\lambda(\bar{x}_n) = \min\left\{1, \left(\frac{0.45}{0.59}\right)^{\frac{1}{n}\sum_{i=1}^{n} x_i} \left(\frac{0.55}{0.41}\right)^{n-\frac{1}{n}\sum_{i=1}^{n} x_i}\right\}$$

$$R = \left\{\min\left\{1, \left(\frac{0.45}{0.59}\right)^{\frac{1}{n}\sum_{i=1}^{n} x_i} \left(\frac{0.55}{0.41}\right)^{n-\frac{1}{n}\sum_{i=1}^{n} x_i}\right\} \le c\right\}$$

← since this will be smaller than 1.

$$= \left\{\left(\frac{0.45}{0.59}\right)^{\frac{1}{n}\sum_{i=1}^{n} x_i} \left(\frac{0.55}{0.41}\right)^{n-\frac{1}{n}\sum_{i=1}^{n} x_i} \le c\right\}$$

$$= \left\{\log\left[\left(\frac{0.45}{0.59}\right)^{\frac{1}{n}\sum_{i=1}^{n} x_i} \left(\frac{0.55}{0.41}\right)^{n-\frac{1}{n}\sum_{i=1}^{n} x_i}\right] \le \log(c)\right\}$$

$$= \left\{\frac{1}{n}\sum_{i=1}^{n} x_i \log\left(\frac{0.45}{0.59}\right) + \left(n - \frac{1}{n}\sum_{i=1}^{n} x_i\right)\log\left(\frac{0.55}{0.41}\right) \le \log(c)\right\}$$

$$= \left\{\bar{x}_n \log\left(\frac{0.45}{0.59}\right) + n\log\left(\frac{0.55}{0.41}\right) - \bar{x}_n \log\left(\frac{0.55}{0.41}\right) \le \log(c)\right\}$$

$$= \left\{\bar{x}_n\left(\log\left(\frac{0.45}{0.59}\right) - \log\left(\frac{0.55}{0.41}\right)\right) \ge \log(c) - n\log\left(\frac{0.55}{0.41}\right)\right\}$$

$$R = \left\{\bar{x}_n \ge \frac{\log(c) - n\log\left(\frac{0.55}{0.41}\right)}{\log\left(\frac{0.45}{0.59}\right) - \log\left(\frac{0.55}{0.41}\right)}\right\}$$

② let $c' = \dfrac{\log(c) - n\log\left(\frac{0.55}{0.41}\right)}{\log\left(\frac{0.45}{0.59}\right) - \log\left(\frac{0.55}{0.41}\right)}$

$R = \{\bar{x}_n \ge c'\}$

$$P(\text{Type I}) = P(R \mid H_0) = P(\bar{x}_n \ge c' \mid p = 0.59)$$

CLT approx. →

$$\approx P\left(\frac{\sqrt{n}(\bar{x}_n - 0.59)}{\sqrt{0.41 * 0.59}} \ge \frac{\sqrt{n}(c' - 0.59)}{\sqrt{0.41 \cdot 0.59}} \,\middle|\, p = 0.59\right)$$

$$= 1 - \Phi\left(\frac{\sqrt{n}\,(c' - 0.59)}{\sqrt{0.41 * 0.59}}\right) \leq \alpha = 0.01$$

$$\Rightarrow \frac{\sqrt{n}\,(c' - 0.59)}{\sqrt{0.41 \cdot 0.59}} \geq z_\alpha = 2.326$$

$$\sqrt{n} \geq \frac{2.326\,(\sqrt{0.41 \cdot 0.59})}{c' - 0.59}$$

$$\sqrt{n} \geq \frac{2.326\,(\sqrt{0.41 \cdot 0.59})}{\left[\frac{\log c - n\log\left(\frac{0.55}{0.41}\right)}{\log\left(\frac{0.45}{0.59}\right) - \log\left(\frac{0.55}{0.41}\right)}\right] - 0.59}$$

$$\sqrt{n} \geq \frac{2.326\,(\sqrt{0.41 \cdot 0.59})}{\frac{\log(c) - n[0.1276]}{-0.2452} - 0.59}$$

$$\sqrt{n} \geq \frac{2.326\,(\sqrt{0.41 \cdot 0.54})}{\frac{\log(c)}{-0.2452} + n(0.5204) - 0.59}$$

$$\geq \frac{2.326\,(\sqrt{0.41 \cdot 0.59})}{-0.59}$$

$$\Rightarrow$$

$$n \geq 3.75968$$

$$P(\text{Type II error}) = 1 - P(R \mid H_1) = P(\bar{X}_n < c' \mid p = 0.45)$$

$$= P\left(\frac{\sqrt{n}\,(\bar{X}_n - 0.45)}{\sqrt{0.45 \cdot 0.55}} < \frac{\sqrt{n}\,(c' - 0.45)}{\sqrt{0.45 \cdot 0.55}} \,\middle|\, p = 0.45\right)$$

$$= \Phi\left[\frac{\sqrt{n}(c' - 0.45)}{\sqrt{0.45 \cdot 0.55}}\right] \leq 0.01$$

bound which occurs when $\Phi(-2.326) = 0.01$

$$\frac{\sqrt{n}(c' - 0.45)}{\sqrt{0.45 \cdot 0.55}} \leq -2.326$$

$$\sqrt{n}(c' - 0.45) \leq -2.326\left(\sqrt{0.45 \cdot 0.55}\right)$$

$$\sqrt{n} \leq \frac{-2.326\left(\sqrt{0.45 \cdot 0.55}\right)}{c' - 0.45}$$

$$\sqrt{n} \leq \frac{-2.326\left(\sqrt{0.45 \cdot 0.55}\right)}{\left[\frac{\log(c) - n\log\left(\frac{0.55}{0.41}\right)}{\log\left(\frac{0.45}{0.59}\right) - \log\left(\frac{0.55}{0.41}\right)}\right] - 0.45}$$

$$\sqrt{n} \leq \frac{-2.326\left(\sqrt{0.45 \cdot 0.55}\right)}{\left[\frac{\log(c) - n(0.1276)}{-0.2452}\right] - 0.45}$$

$$\downarrow$$

$$\leq \frac{-2.326\left(\sqrt{0.45 \cdot 0.55}\right)}{-0.45}$$

$$\Rightarrow \sqrt{n} \geq 2.571489754$$

$$n \geq 6.612559556$$

Hence, approximate sample size for both type I and type II error to be at most 0.01 is 6.613.

6] ① $X_1, \dots, X_n$ are independently sampled from Pareto(θ, v)

$$f(x|\theta) = \begin{cases} \frac{\theta v^\theta}{x^{\theta+1}} & \text{if } x \geq v \quad \theta > 0, v > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$L(\theta, v) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{\theta v^\theta}{x_i^{\theta+1}}$$

$$l(\theta, v) = \log L(\theta, v) = \log\left(\prod_{i=1}^{n} \frac{\theta v^\theta}{x_i^{\theta+1}}\right) = \sum_{i=1}^{n} \log\left(\frac{\theta v^\theta}{x_i^{\theta+1}}\right)$$

$$= \sum_{i=1}^{n} \log(\theta v^\theta) - \log(x_i^{\theta+1}) = \sum_{i=1}^{n} \log(\theta v^\theta) - (\theta+1)\log(x_i)$$

$$= \sum_{i=1}^{n} \log(\theta) + \theta\log(v) - \theta\log(x_i) - \log(x_i)$$

$$l(\theta, v) = n\log(\theta) + n\theta\log(v) - \theta\sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\log(x_i)$$

$$\frac{\partial l}{\partial v} = \frac{n\theta}{v} = 0 \Rightarrow \hat{v}_{MLE} = 0 \leftarrow \text{however can't be case}$$

since by assumption $v > 0$. We must find a different argument.

$f(x|\theta) = \frac{\theta v^\theta}{x^{\theta+1}}$ is an increasing function with respect to $v \to$ maximized w/t v when v is at greatest possible value.

this pdf is defined for $x \geq v$, which means for $x_1, \dots, x_n$ $\min\{x_1, \dots, x_n\} \geq v$. Hence, maximum possible value of v is $v = \min\{x_1, \dots, x_n\}$, and thus $\boxed{\hat{v}_{MLE} = \min\{x_1, \dots, x_n\}}$.

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + n\log(\nu) - \sum_{i=1}^{n} \log(X_i) = 0$$

$$\frac{n}{\theta} = \sum_{i=1}^{n} \log(x_i) - n\log(\nu)$$

$$\hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} \log(X_i) - n\log(\nu)}$$

② $H_0: \theta = 1$, ν unknown

$H_1: \theta \neq 1$, ν unknown

$$\lambda(\bar{X}_n) = \frac{\sup_{\Theta_0} L(\theta \mid X_n)}{\sup_{\Theta} L(\theta \mid X_n)} = \frac{\sup_{\Theta_0} L(\theta = 1 \mid X_n)}{\sup_{\Theta} L(\theta \neq 1 \mid X_n)}$$

corresponds to sample mean i.e. $\hat{\theta}_{MLE}$

$$= \frac{\prod_{i=1}^{n} \frac{(1)\,\nu^{(1)}}{X_i^{(1+1)}}}{\prod_{i=1}^{n} \frac{(\bar{\theta}_n)\,\nu^{(\bar{\theta}_n)}}{X_i^{(\bar{\theta}_n+1)}}} = \frac{\frac{n\nu}{\prod_{i=1}^{n} X_i^2}}{\frac{n\bar{\theta}_n \nu^{\bar{\theta}_n}}{\prod_{i=1}^{n} X_i^{(\bar{\theta}_n+1)}}} = \frac{\frac{\nu}{\prod_{i=1}^{n} X_i^2}}{\frac{\bar{\theta}_n \nu^{\bar{\theta}_n}}{\prod_{i=1}^{n} X_i^{\bar{\theta}_n} X_i}}$$

$$= \frac{\frac{\nu}{\prod_{i=1}^{n} X_i}}{\frac{\bar{\theta}_n \nu^{\bar{\theta}_n}}{\prod_{i=1}^{n} X_i^{\bar{\theta}_n}}}$$

$$\lambda(X_n) = \frac{\sup_{\theta_0} L(\theta=1 | X_n)}{\sup_{\theta} L(\underset{\hat{\theta}_{MLE}}{\underset{\wedge}{\theta}} \neq 1 | X_n)} = \frac{\prod_{i=1}^{n} \frac{(1)V^1}{x_i^2}}{\prod_{i=1}^{n} \frac{\theta V^\theta}{x_i^{\theta+1}}}$$

$$= \frac{\frac{nV}{\prod_{i=1}^{n} x_i^{-2}}}{\frac{\theta^n V^{\theta n}}{\prod_{i=1}^{n} x_i^{\theta+1}}} = \left(\frac{nV}{\prod_{i=1}^{n} x_i^{-2}}\right)\left(\frac{\prod_{i=1}^{n} x_i^{\theta+1}}{\theta^n V^{\theta n}}\right) = \left(\prod_{i=1}^{n} \frac{x_i^{\theta+1}}{x_i^{-2}}\right)\left(\frac{nV}{\theta^n V^{\theta n}}\right)$$

$$\lambda(X_n) = \left(\prod_{i=1}^{n} x_i^{\theta-1}\right)\left(\frac{nV}{\theta^n V^{\theta n}}\right)$$

since we have a 2-sided hypothesis:

$$R = \{\lambda(x_n) \leq \underbrace{c_1' \text{ or } \lambda(X_n) \geq c_2'}_{\text{constants}}\}$$

$\lambda(X_n) \leq c_1'$:

$$\frac{1}{\theta}\left(\prod_{i=1}^{n} x_i^{\theta-1}\right)(V^{1-\theta}) \leq c_1'$$

$$\log\left(\frac{1}{\theta}\right) + \log(V^{1-\theta}) + \log\left(\prod_{i=1}^{n} x_i^{\theta-1}\right) \leq \log(c_1)$$

$$\log\left(\prod_{i=1}^{n} x_i^{\theta-1}\right) \leq \underbrace{\log(c_1) - \log\left(\frac{1}{\theta}\right) - \log(V^{1-\theta})}_{c_1}$$

$$(\theta-1)\log\left(\prod_{i=1}^{n} x_i\right) \leq c_1$$

$$\theta \log\left(\prod_{i=1}^{n} x_i\right) - \log\left(\prod_{i=1}^{n} x_i\right) \leq c_1$$

$$\left(\prod_{i=1}^{n} x_i^{\theta-1}\right)\left(\frac{nV}{\theta^n V^{\theta n}}\right) \leq c_1'$$

$$\log\left(\prod_{i=1}^{n} x_i^{\theta-1}\right) + \log\left(\frac{nV}{\theta^n V^{\theta n}}\right) \le \log(C_1')$$

$$(\theta-1)\log\left(\prod_{i=1}^{n} x_i\right) + \log(nV) - \log(\theta^n) - \log(V^{\theta n}) \le \log(C_1')$$

$$(\theta-1)\log\left(\prod_{i=1}^{n} x_i\right) - \log(\theta^n) - \theta n \log(V) \le \log(C_1') - \log(nV)$$

$$\underbrace{\log\left(\frac{\prod_{i=1}^{n} x_i}{\theta^n}\right)}_{T(X_n)} \le \underbrace{\frac{\log(C_1') - \log(nV) + \theta n \log(V)}{\theta - 1}}_{C_1}$$

where $\theta = X_{(1)}$ which is assumed under H_1. (where $\hat{\theta}_{MLE}$ exists)

$$\lambda(X_n) \ge C_2'$$

$$\left(\prod_{i=1}^{n} x_i^{\theta-1}\right)\left(\frac{nV}{\theta^n V^{\theta n}}\right) \ge C_2'$$

$$(\theta-1)\log\left(\prod_{i=1}^{n} x_i\right) + \log(nV) - \log(\theta^n) - \log(V^{\theta n}) \ge C_2'$$

$$\underbrace{\log\left(\frac{\prod_{i=1}^{n} x_i}{\theta^n}\right)}_{T(X_n)} \ge \underbrace{\frac{\log(C_2') + \theta n \log(V) - \log(nV)}{\theta - 1}}_{C_2}$$

where $\theta = X_{(1)}$

$H_0: \theta = \mu$
$H_1: \theta \neq \mu$

7] (1) $X_1, \ldots, X_n \overset{iid}{\sim} beta(\mu, 1)$ $\quad X \perp\!\!\!\perp Y$
$Y_1, \ldots, Y_m \overset{iid}{\sim} beta(\theta, 1)$

$$\lambda(X_n, Y_m) = \frac{\sup_{H_0} L(\mu, \theta \mid X_n, Y_m)}{\sup_{H_1} L(\mu, \theta \mid X_n, Y_m)}$$

$$L(\mu, \theta) = \prod_{i=1}^{n} f(X_i \mid \mu) \prod_{j=1}^{m} f(Y_j \mid \theta)$$

$$= \left[\prod_{i=1}^{n} \mu X_i^{\mu-1}\right]\left[\prod_{j=1}^{m} \theta Y_j^{\theta-1}\right]$$

$$= \mu^n \left(\prod_{i=1}^{n} X_i\right)^{\mu-1} \theta^j \left[\prod_{j=1}^{m} Y_j\right]^{\theta-1}$$

under $H_0: \theta = \mu$

$$L(\mu, \theta) = \theta^{j+n} \left(\prod_{i=1}^{n} X_i \prod_{j=1}^{n} Y_j\right)^{\theta-1}$$

Hence, $$\lambda(X_n, Y_m) = \frac{\hat{\theta}_0^{j+n} \left(\prod_{i=1}^{n} X_i \prod_{j=1}^{n} Y_j\right)^{\hat{\theta}_0 - 1}}{\hat{\mu}^n \hat{\theta}^j \left(\prod_{i=1}^{n} X_i\right)^{\hat{\mu}-1} \left(\prod_{j=1}^{m} Y_j\right)^{\hat{\theta}-1}}$$

$$\lambda(X_n, Y_m) = \left(\frac{\hat{\theta}_0^{j+n}}{\hat{\mu}^n \hat{\theta}^j}\right)\left(\prod_{i=1}^{n} X_i\right)^{\hat{\theta}_0 - \hat{\mu}} \left(\prod_{j=1}^{m} Y_j\right)^{\hat{\theta}_0 - \hat{\theta}}$$

For the 2nd part of this question, I'm guessing it's needed to solve for $\hat{\theta}_0$, $\hat{\mu}$ and $\hat{\theta}$....

For θ_0: $l(\theta) = \log L(\theta) = \log\left(\theta^{j+n}\left(\prod_{i=1}^{n} x_i \prod_{j=1}^{n} y_j\right)^{\theta-1}\right)$

$$= (j+n)\log(\theta) + (\theta-1)\log\left(\prod_{i=1}^{n} x_i \prod_{j=1}^{n} y_j\right)$$

$$= (j+n)\log(\theta) + \theta\sum_{i=1}^{n}\log(x_i) + \theta\sum_{j=1}^{n}\log(y_j) - 1\log\left(\prod_{i=1}^{n} x_i \prod_{j=1}^{n} y_j\right)$$

$$\frac{\partial l(\theta)}{\partial \theta} = \frac{j+n}{\theta} + \sum_{i=1}^{n}\log(x_i) + \sum_{j=1}^{n}\log(y_j) = 0$$

$$\hat{\theta}_{0_{MLE}} = \frac{-\,j+n}{\sum_{i=1}^{n}\log(x_i) + \sum_{j=1}^{n}\log(y_j)}$$

For $\hat{\theta}, \hat{\mu}$: $l(\theta,\mu) = \log\left(\mu^n\left(\prod_{i=1}^{n} x_i\right)^{\mu-1}\theta^j\left(\prod_{j=1}^{n} y_j\right)^{\theta-1}\right)$

$$l(\theta,\mu) = n\log(\mu) + (\mu-1)\sum_{i=1}^{n}\log(x_i) + j\log(\theta) + (\theta-1)\sum_{j=1}^{n}\log(y_j)$$

$$\frac{\partial l}{\partial \theta} = \frac{j}{\theta} + \sum_{j=1}^{n}\log(y_j) = 0 \qquad \hat{\theta}_{MLE} = \frac{-j}{\sum_{j=1}^{n}\log(y_j)}$$

$$\frac{\partial l}{\partial \mu} = \frac{n}{\mu} + \sum_{i=1}^{n}\log(x_i) = 0 \qquad \hat{\mu}_{MLE} = \frac{-n}{\sum_{i=1}^{n}\log(x_i)}$$

$$\lambda(X_n, Y_m) = \frac{\hat{\theta}_0^{\,j}\,\hat{\theta}_0^{\,n}}{\hat{\mu}^n\,\hat{\theta}^{\,j}}\underbrace{\left(\prod_{i=1}^{n} x_i\right)^{\hat{\theta}_0-\hat{\mu}}\left(\prod_{j=1}^{m} y_j\right)^{\hat{\theta}_0-\hat{\theta}}}_{=1}$$

$$= \left(\frac{j+n}{j}\right)^j\left(\frac{j+n}{n}\right)^n\left(1-\frac{\sum_{i=1}^{n}\log x_i}{\sum_{i=1}^{n}\log x_i + \sum_{j=1}^{m}\log y_j}\right)^j\left(\frac{\sum_{i=1}^{n}\log x_i}{\sum_{i=1}^{n}\log x_i + \sum_{i=1}^{m}\log y_j}\right)^n$$

$$R = \{\lambda(X_n, Y_m) \leq c\}\ \cancel{\text{or } \lambda(X_n, Y_m) \geq c_2\}}$$

↓ motivated by week 3 discussion thread

$\cancel{\lambda(x_n, y_m) \le c_1}$

$\cancel{\lambda = \frac{\sum_{i=1}^{n} \log x_i}{\sum_{i=1}^{n} \log x_i + \sum_{j=1}^{m} \log y_j}}$

let $T = \frac{\sum_{i=1}^{n} \log x_i}{\sum_{i=1}^{n} \log x_i + \sum_{j=1}^{m} \log y_j}$

Then, $\lambda(x_n, y_m)$ is a unimodal function of T,

$\Rightarrow R = \{T \le c_1 \text{ or } T \ge c_2\}$

is equivalent rejection region.

② $T = \frac{\sum_{i=1}^{n} \log x_i}{\sum_{i=1}^{n} \log x_i + \sum_{j=1}^{m} \log y_j} = \frac{\sum_{i=1}^{n} -\log x_i}{\sum_{i=1}^{n} -\log x_i + \sum_{j=1}^{m} -\log y_j}$

using hints in question: $-\log x_i \sim Gamma(1, \frac{1}{\mu})$

since x_i is iid; $W = \sum -\log x_i \sim Gamma(n, \frac{1}{\mu})$

$-\log y_j \sim Gamma(1, \frac{1}{\theta})$; y_j is iid; $V = \sum -\log y_j$

$\sim Gamma(m, \frac{1}{\theta})$; since under H_0, $\theta = \mu$

$V \sim Gamma(m, \frac{1}{\mu})$; by question assumption, X's are independent of Y, hence W is independent of V.

Thus, $T = \frac{W}{(W+V)} \sim Beta(n, m)$

since $R = \{T \le c_1 \text{ or } T \le c_2\}$ for c_1 and c_2 constants,

This means $P(T \le c_1) + P(T \ge c_2) = \alpha$ and

①→

$P(T \le c_1) = P(T \ge c_2)$

② $(1-c_1)^m c_1^n = (1-c_2)^m c_2^n$

where $\alpha = 0.05$, and c_1 & c_2 are determined by solving the system of equations ① & ②

③ $T(X_{13}, Y_{17}) = 0.6086678$ ← see attached code

$R = P(T \le c_1' \text{ or } T \ge c_2' \mid H_0)$

$= 1 - [P(T \le c_1' \mid H_0) - P(T \ge c_2' \mid H_0)]$

$= 1 - [P(T \le 0.608 \mid H_0) - P(T \ge 0.608 \mid H_0)]$

where H_0 assumes $T \sim Beta(13, 17)$

$= 1 - 2P(T \le 0.608 \mid H_0)$

$= 1 -$

p-value: $1 - P(T \le 0.608 \mid H_0)$ where H_0 assumes $T \sim Beta(13, 17)$

$= 1 - 0.973205$ ← see code for computation.

$\boxed{\text{p-value} = 0.026795}$

Yes, I reject the null hypothesis at 5% significance level $\alpha = 0.05$.

8] ① sample mean: 89.55475
sample SD: 14.90353 ← from attached code

② $\alpha = 0.01$; $H_0: \mu = \mu_0$ $\mu_0 = 100$
$H_1: \mu < \mu_0$

From LECTURE 9, Slide 20: $\Leftrightarrow$

$$R = \left\{ \frac{\sqrt{n}(\bar{x}_n - \mu_0)}{s} \leq -t_{n-1}(\alpha) \right\}$$

$$R = \left(\frac{\sqrt{n}(\bar{x}_n - \mu_0)}{s} \leq -t_{19}(0.01) \right)$$

$t_{19}(0.01) = -2.539483$

$$P\left(\frac{\sqrt{n}(\bar{x}_n - \mu_0)}{s} \leq -t_{19}(0.01) \middle| H_0 \right)$$

where H_0 assumes

$$= P\left(\frac{\sqrt{n}(\bar{x}_n - \mu_0)}{s} \leq 2.539483 \middle| H_0 \right) \sim N(100, \sigma^2)$$

$= 0.071835$

Hence, we fail to reject null hypothesis. The data does not indicate mean DL reading for smokers is significantly lower than 100.

$\frac{\sigma^2}{n} = s^2$

$\sigma^2 = s^2 n$

$\sigma = s\sqrt{n}$

$\sigma = 14.90353(\sqrt{20})$

$\sigma^2 = 4442.30479$

9] $\lambda(x) = \dfrac{L(H_0 | x)}{L(H_1 | x)} = \begin{cases} \frac{0.01}{0.06} & x=1 \\ \frac{0.01}{0.05} & x=2 \\ \frac{0.01}{0.04} & x=3 \\ \frac{0.01}{0.03} & x=4 \\ \frac{0.02}{0.02} & x=5 \\ \frac{0.01}{0.01} & x=6 \\ \frac{0.93}{0.79} & x=7 \end{cases}$

$R = \{\lambda(x) \leq c\} = \begin{cases} \{1,2,3,4,5,6,7\} & \text{if } c \geq 1.177 \\ \{1,2,3,4,5,6\} & \text{if } 1 \leq c < 1.177 \\ \{1,2,3,4\} & \text{if } \frac{1}{3} \leq c < 1 \\ \{1,2,3\} & \text{if } \frac{1}{4} \leq c < \frac{1}{3} \\ \{1,2\} & \text{if } \frac{1}{5} \leq c < \frac{1}{4} \\ \{1\} & \text{if } \frac{1}{6} \leq c < \frac{1}{5} \\ \emptyset & \text{if } 0 \leq c < \frac{1}{6} \end{cases}$

let $\alpha = 0.04$. By Neyman-Pearson lemma,

$P(\text{Type I}) = \alpha$ has max power for tests

$P(\lambda(X_n) \leq c \mid H_0) = \alpha$

$\Rightarrow P(X = 1, 2, 3, 4 \mid H_0) = 0.04$

or

$\Rightarrow$ LRT should be

where $\mathbb{I}$ is indicator function.

$$\lambda = \mathbb{I}(\lambda(x) \leq c)$$

$$= \mathbb{I}\{x = 1 \text{ or } 2 \text{ or } 3 \text{ or } 4\}$$

$$P(\text{Type II error}) = P(\lambda(x_n) > c \mid H_1)$$

$$= 1 - P(\lambda(x_n) \leq c \mid H_1)$$

$$= 1 - P(\mathbb{I}\{x = 1 \text{ or } 2 \text{ or } 3 \text{ or } 4\} \mid H_1)$$

$$= 1 - (0.06 + 0.05 + 0.04 + 0.03)$$

$$= 0.82$$

10] ① $f(X|\theta) = \theta x^{\theta - 1}, \ 0 < x < 1$

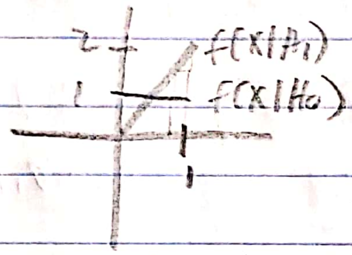
$\alpha = 0.05, \ H_0: \theta = 1, \ H_1: \theta = 2$

We want the LRT w/ P(Type I error) = α for maximum power test.

$$\lambda(\bar{X}_n) = \frac{\sup_{H_0} L(\theta|\bar{X}_n)}{\sup_{H_1} L(\theta|\bar{X}_n)} = \frac{L(\theta = 1|\bar{X}_n)}{L(\theta = 2|\bar{X}_n)} \leq c$$

$f(X|H_0) = 1x^{1-1} = 1, \ 0 < x < 1$

$f(X|H_1) = 2x^{2-1} = 2x, \ 0 < x < 1$



$$\lambda(\bar{X}_n) = \frac{1}{2(\bar{X}_n)} \leq c$$

For significance level, $\alpha = 0.05$

$$P(\bar{X}_n \geq c' | \theta = 1) = 0.05$$

from graph, $f(X|H_0) > f(X|H_1)$ for $x \in [0, 0.5)$

$f(X|H_1) > f(X|X_0)$ $x \in [0.5, 1]$

$$R = \{\lambda(X_n) \leq c\} \Rightarrow \left\{\frac{1}{2(\bar{X}_n)} \leq c\right\}$$

$$= \left\{\bar{X}_n \leq \frac{1}{2c}\right\}$$

$$P\left(\bar{X}_n \leq \frac{1}{2c} \Big| \theta = 1\right) = 0.05$$

$$\Rightarrow \frac{1}{2c} = 0.05, \ c = 10$$

we have found c' such that $P(\bar{X}_n \geq c' | H_0) = 0.05$
$\downarrow 0.95$ since $H_0 \sim$ uniform $(0,1)$

② $B(\theta) = P(X_n \in R | \theta)$ ← power function
power of test $= 1 - P(X_n \notin R | \theta_1)$
$= P(X_n \in R | \theta_1)$
$= P(R | \theta = 2)$
where $R = \{\bar{X}_n \geq c'\}$ where $c' = 0.95$

$P(\bar{X}_n \geq 0.95 | \theta = 2)$
$= \dfrac{0.05(1.9) + \frac{1}{2}(0.05)(0.1)}{2}$

power of test $= 0.04875$

③ $P(\bar{X}_n \geq 0.8 | H_0) = 1 - 0.8 = 0.2$
p-value is 0.2

④ $H_1 : \theta > 1$ ← correction in week 3 discussion

Yes. The test is uniformly most powerful against alternative hypothesis $H_1 : \alpha > 1$. From slide 23 of lecture 8, this is an implication of Neyman-Pearson lemma. [Since this is one-sided hypothesis test and we found LRT w/ significance level α.]