Dec 1st,

HW3 & Mid 2.

***9.50** Let $Y_1, Y_2, \dots, Y_n$ denote a random sample from the uniform distribution over the interval (θ_1, θ_2). Show that $Y_{(1)} = \min(Y_1, Y_2, \dots, Y_n)$ and $Y_{(n)} = \max(Y_1, Y_2, \dots, Y_n)$ are jointly sufficient for θ_1 and θ_2.

9.50 $P(Y_1, \dots, Y_n \mid \theta_1, \theta_2) = \left(\frac{1}{\theta_2 - \theta_1}\right)^n 1\{\theta_1 \le Y_1, \dots, Y_n \le \theta_2\}$

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

So $Y_{(1)}, Y_{(n)}$ are SS.

1

Let $\{X_n\}_n$ and $\{Y_n\}_n$ be two sequences of random variables. Suppose $X_n \to_p x$ and $Y_n \to_p y$, as $n \to \infty$, where $x, y \in \mathbb{R}$ are constants. In the class, I proved $X_n + Y_n \to_p x + y$. Please give a proof of the result: $X_n Y_n \to_p xy$. You should prove it from the definition of convergence in probability (i.e. you can't use theorem 9.2).

(Hint: $|X_n Y_n - xy| = |X_n Y_n - xY_n + xY_n - xy| \le \underbrace{|X_n Y_n - xY_n|} + \underbrace{|xY_n - xy|}$)

1. Solution 1:

For $X_n \xrightarrow{P} x$, $Y_n \xrightarrow{P} y$, so for any $\varepsilon > 0$, $\exists N > 0$,

s.t. $\forall n > N$, we have $P(|X_n - x| > \varepsilon) \le \varepsilon$, $P(|Y_n - y| > \varepsilon) \le \varepsilon$.

So $\forall n > N$, $|X_nY_n - xy| \le |X_nY_n - xY_n| + |xY_n - xy|$

$$\le |X_n - x|\,|Y_n| + |x|\,|Y_n - y|$$

$$\le \varepsilon(|y| + \varepsilon) + \varepsilon|x|$$

$$\le \varepsilon M. \quad (M \text{ is a constant}).$$

holds with prob at least $1-2\varepsilon$.

Therefore, $X_nY_n \xrightarrow{P} xy$.

$P(|X_n^2 - X^2| > \varepsilon) \to 0$.

by definition

Solution 2:

Verify ~~that~~ if $X_n \xrightarrow{P} X$, then $X_n^2 \xrightarrow{P} X^2$. ①

So $(X_n+Y_n)^2 \xrightarrow{P} (X+Y)^2$, $(X_n-Y_n)^2 \xrightarrow{P} (X-Y)^2$,

hence $X_nY_n = \dfrac{(X_n+Y_n)^2 - (X_n-Y_n)^2}{4} \xrightarrow{P} \dfrac{(X+Y)^2 - (X-Y)^2}{4} = XY$.

$X_n + Y_n \xrightarrow{P} X+Y$ ②

STAT 4102 Midterm 2

2020 Fall

1 (25 points)

Let $X_1, ..., X_n$ be i.i.d. random variables with pdf

$$f(x) = \theta x^{\theta-1}, x \in (0,1), \theta \in (0,\infty).$$

1. What is the likelihood function? (4 points)
2. Find a one-dimensional sufficient statistic for θ. Note that the vector $(X_1, ..., X_n)$ is sufficient by definition, but it is n-dimensional. (7 points)
3. Find the method of moment estimator for θ. (7 points)
4. It is consistent? Prove your result. (7 points) $=1$

$$EX_1 = \int_0^1 x\,\theta x^{\theta-1}\,dx = \int_0^1 \theta x^{\theta}\,dx = \theta \frac{x^{\theta+1}}{\theta+1}\Big|_0^1 = \frac{\theta}{\theta+1}$$

Solution:

$g(\prod_i x_i) \cdot h(x_1, \dots, x_n)$

1. $L(\theta|x_1, ..., x_n) = \theta^n(\prod_i x_i)^{\theta-1}$.

$T(X_1, \dots, X_n)$

2. By the factorization theorem, $\prod_i X_i$ is a sufficient statistic.

3. $EX_1 = \frac{\theta}{\theta+1}$, and thus a method of moments estimator is

$$\frac{\hat\theta}{\hat\theta+1} = \bar X \Rightarrow \qquad \hat\theta = \frac{\bar X}{1-\bar X}.$$

4. It is consistent. By the law of large number, $\bar X \to_p EX_1$. Then by theorem 9.2, we get $\frac{\bar X}{1-\bar X} \to_p \theta$, since $f(x) = \frac{x}{1-x}$ is continuous on $(0,1)$.

$$\hat\theta = \frac{\bar X}{1-\bar X} \xrightarrow{P} \frac{EX_1}{1-EX_1} = \theta.$$

2 (25 points)

Let $X_1, ..., X_n$ be i.i.d. random variables with pdf

$$f(x) = \frac{2x}{\theta^2}, x \in (0,\theta), \theta \in (0,\infty).$$

1. What is the likelihood function? (4 points)
2. Find the MLE for θ. (10 points)
3. Is it consistent? Prove your result. (6 points)
4. Find the MLE for $\exp(\theta^2)$. Is it consistent? (5 points)

MLE for e^{θ^2} is $e^{\hat\theta^2}$

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

$L(\theta|X_1,\cdots,X_n) = \frac{2}{\theta^2}\prod_{i=1} X_i \cdot 1\{0<X_1,\cdots,X_n<\theta\}.$

$= \frac{2}{\theta^2}\prod_{i=1}^{n} X_i \cdot 1\{X_{(1)}>0\}\, 1\{X_{(n)}\le\theta\}$

Solution:

1. $L(\theta|x_1,...,x_n) = \frac{2^n}{\theta^{2n}}\prod_i x_i 1_{x_i\le\theta}1_{x_i>0} = \frac{2^n}{\theta^{2n}}1_{\min x_i>0}1_{\max x_i\le\theta}\prod_i x_i$.

2. $L(\theta) = (\prod_i x_i 1_{\min x_i>0}2^n)(1_{\max x_i\le\theta}\frac{1}{\theta^{2n}})$. Since the first factor of the right hand side is a positive constant, the second factor is a decreasing function on $[\max x_i, \infty)$, and it is 0 when $\theta < \max x_i$, we know the likelihood function is maximized at $\max x_i$. Thus $\hat{\theta} = \max X_i$.

3. Let $\epsilon > 0$. Because $\max X_i \le \theta$, we have

$$P(|\max X_i - \theta| > \epsilon) = P(\max X_i < \theta - \epsilon).$$

$\boxed{P(|X_{(n)}-\theta|>\varepsilon)\to 0.}$

$= P(\theta - X_{(n)} > \varepsilon)$

Now, let $F(x)$ be the c.d.f. of X_1. We know that $F(x) < 1$, for any $x < \theta$. Since

$$P(\max X_i < \theta - \epsilon) = P(X_1, X_2, ..., X_n < \theta - \epsilon) = P(X_1 \le \theta - \epsilon)^n = F(\theta - \epsilon)^n,$$

$= P(X_{(n)} < \theta - \varepsilon)$

we get

$$P(|\max X_i - \theta| > \epsilon) \to 0, \text{ as } n \to \infty.$$

4. By the invariance property of MLE, the MLE of $\exp(\theta^2)$ is $\exp\{(\max X_i)^2\}$. We've shown that $\max X_i \to_p \theta$. By theorem 9.2, the MLE is again consistent.

3 (25 points)

Let $X_1, ..., X_n$ be i.i.d. Gamma(α, β) distributed and $\alpha > 0$ is known. The p.d.f. is

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}\exp(-x/\beta), x > 0.$$

1. Find the MLE of β. (5 points)
2. Suppose we want to test $H_0 : \beta = \beta_0$ against $H_1 : \beta \ne \beta_0$, for some known and fixed $\beta_0 > 0$. Develop a level α test for it, using the MLE as the test statistic. (15 points)
3. What is the power of your test at the point $\beta_1 \ne \beta_0$? (5 points)

You may use the fact that $\sum_i^n X_i \sim$ Gamma$(n\alpha, \beta)$.

Solution:

1. The log-likelihood function is

$$l(\beta|x_1, ..., x_n) = -n\log(\Gamma(\alpha)) - n\alpha\log(\beta) + (\alpha - 1)(\sum_i \log(x_i)) - \frac{1}{\beta}\sum_i x_i.$$

Taking a derivative w.r.t. β, we have

$$-n\alpha\frac{1}{\beta} + \frac{1}{\beta^2}\sum_i x_i.$$

By setting it to 0, we get $\frac{\sum X_i}{n\alpha}$. Because the second derivative

$$n\alpha\frac{1}{\beta^2} - \frac{2}{\beta^3}\sum_i x_i$$

$H_0: \beta = \beta_0$ v.s. $H_1: \beta \neq \beta_0$ $X_1, \dots, X_n \overset{i.i.d.}{\sim} Gamma(\alpha, \beta)$.

$\hat{\beta} = \frac{\sum X_i}{n\alpha} = \frac{\bar{X}}{\alpha}$

is negative at this point, we know that the MLE is

$$\hat{\beta} = \frac{\sum X_i}{n\alpha}.$$

$n\alpha\hat{\beta} = \sum X_i \overset{H_0}{\sim} Gamma(n\alpha, \beta_0)$

2. Let $\gamma_{1-\omega/2}$ and $\gamma_{\omega/2}$ be two quantiles of the Gamma$(n\alpha, \beta_0)$ distribution. Then we have

$$P_{H_0}(\gamma_{1-\omega/2}/(n\alpha) \leq \hat{\beta} \leq \gamma_{\omega/2}/(n\alpha)) = \omega.$$

w-level

Thus we can reject the null when $\hat{\beta} \leq \gamma_{1-\omega/2}/(n\alpha)$ or $\hat{\beta} \geq \gamma_{\omega/2}/(n\alpha)$, and this is a level ω test.

3. The power is

$P_{\beta_1}(\hat{\beta} \geq \gamma_{\frac{w}{2}}$ or $\hat{\beta} \leq \gamma_{1-\frac{w}{2}})$.

$$P(\text{Gamma}(n\alpha, \beta_1) \geq \gamma_{\omega/2}) + P(\text{Gamma}(n\alpha, \beta_1) \leq \gamma_{1-\omega/2}).$$

$n\alpha\hat{\beta} = \sum X_i \overset{\beta=\beta_1}{\sim} Gamma(n\alpha, \beta_1)$

4 (25 points)

A coin operated soda machine was designed to discharge an average of 7 ounces of soda per cup. In a test of the machine, 30 cups of beverage were drawn from the machine and measured. The mean and standard deviation of the measurements were 7.1 ounces and 0.12 ounce, respectively.

✗ t-test

1. Do these data present sufficient evidence to indicate that the mean discharge differs from 7 ounces? Use $\alpha = 0.10$. If you use a formula for this problem, please verify every condition needed for it. (20 points)

Z-test

2. What is the p-value? Write your answer in terms of the c.d.f of normal distribution. You do not need to calculate the number. (5 points)

(Some quantiles you may need: $z_{0.25} = 1.96$, $z_{0.05} = 1.64$, $z_{0.1} = 1.28$)

Solution:

1. Let $\bar{X}$ be the sample mean, and S^2 be the sample variance. We use the statistic

$$Z = \frac{\bar{X} - 7}{\sqrt{S}/\sqrt{n}}$$

for testing $H_0 : \mu = 7$ vs $H_1 : \mu \neq 7$. If we assume that the volume of beverage in the 30 cups $X_1, ..., X_n$ are i.i.d., with mean μ and variance σ^2, then under the null hypothesis, by the central limit theorem,

$$\bar{X} - \mu \sim_{approx} N(0, \sigma^2/n).$$

Since S^2 is a consistent estimator for σ^2, we have that under the null hypothesis,

$$Z = \frac{\bar{X} - 7}{\sqrt{S}/\sqrt{n}} \sim_{approx} N(0, 1).$$

Let z_α denote the corresponding quantile of the standard normal distribution. Then we know a test that rejects the null when $|Z| \geq z_{\alpha/2}$ is a level α test. Using the data provided, we have $Z = 4.564$ and $z_{0.05} = 1.64$. Thus we reject the null hypothesis.

2. Let ϕ be the c.d.f. of the standard normal distribution. The p-value is $2\phi(-|Z|)$.

$\phi(x) = P(Z \leq x)$, $Z \sim N(0,1)$.