HW5

Problem 1.

Normal Distribution with unknow μ and known σ. Show that $\bar{Y}$ is a sufficient statistic for μ.

If the likelihood can be factored into two nonnegative functions, then it is a sufficient statistic

$$L(y_1, y_2, \dots, y_n|\theta) = g(u,\theta) * h(y_1, y_2, \dots, y_n)$$

Then

$$L(y_1, y_2, \dots, y_n|\theta) = L(y_1|\theta) * L(y_2|\theta) * \dots * L(y_n|\theta)$$

$$L(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2}$$

$$\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{\sum_{i=1}^{n}\left(-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2\right)} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2} =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2}$$

Let's resolve part of the exponent:

$$\sum_{i=1}^{n}(y_i-\mu)^2 = \sum_{i=1}^{n}(y_i^2 - 2*y_i*\mu + \mu^2) = \sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} 2*y_i*\mu + \sum_{i=1}^{n} \mu^2 =$$

$$= \sum_{i=1}^{n} y_i^2 - 2*\mu*\sum_{i=1}^{n} y_i + n*\mu^2 = \sum_{i=1}^{n} y_i^2 - 2*\mu*\frac{n}{n}\sum_{i=1}^{n} y_i + n*\mu^2 =$$

$$\sum_{i=1}^{n} y_i^2 - 2*\mu*n\bar{y} + n*\mu^2$$

Let's substitute back in the equation:

$$\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2} = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2 - 2*\mu*n\bar{y} + n*\mu^2\right)} =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2\right)} * e^{-\frac{1}{2\sigma^2}\left(-2*\mu*n\bar{y} + n*\mu^2\right)} =$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2\right)} * e^{\frac{n}{2\sigma^2}\left(2\mu\bar{y} - \mu^2\right)}$$

Then the factorization is:

$$L(y_1, y_2, \dots, y_n|\theta) = g(\bar{y},\theta) * h(y_1, y_2, \dots, y_n)$$

$$g(\bar{y},\theta) = e^{\frac{n}{2\sigma^2}(2\mu\bar{y}-\mu^2)}$$

$$h(y_1, y_2, \dots, y_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n * e^{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} y_i^2\right)}$$

Since it satisfies the factorization, we can say that $\bar{Y}$ is a sufficient statistic for μ.

Alejandro C Parra Garcia

Problem 2.

$$f(y) = \alpha\beta^{\alpha}y^{-(\alpha+1)}, if\ y \geq \beta$$

$$L(y|\theta) = \theta\beta^{\theta}y^{-(\theta+1)}$$

$$L(y_1, y_2, \dots, y_n|\theta) = g(u, \theta) * h(y_1, y_2, \dots, y_n), u = \prod_{i=1}^{n} y_i$$

$$L(y_1, y_2, \dots, y_n|\theta) = \prod_{i=1}^{n} \theta\beta^{\theta}y_i^{-(\theta+1)}$$

$$\prod_{i=1}^{n} \theta\beta^{\theta}y_i^{-(\theta+1)} = \left(\theta\beta^{\theta}\right)^n * \prod_{i=1}^{n} y_i^{-(\theta+1)} = \left(\theta\beta^{\theta}\right)^n * \prod_{i=1}^{n} y_i^{-(\theta+1)}$$

Problem 3.

a) Use Method of moments

Since we are resolving for only one parameter, we use only the 1st moment:

$$E(x) = \int_0^1 x^1 * (\theta+1)x^{\theta}\,dx$$

$$\int_0^1 x * (\theta+1)x^{\theta}\,dx = \int_0^1 (\theta+1)x^{\theta+1}\,dx = \left[(\theta+1)\frac{x^{\theta+2}}{\theta+2}\right]_0^1 =$$

$$\left[(\theta+1)\frac{x^{\theta+2}}{\theta+2}\right]_0^1 = \left((\theta+1)\frac{1^{\theta+2}}{\theta+2}\right) - \left((\theta+1)\frac{0^{\theta+2}}{\theta+2}\right) =$$

$$= \left((\theta+1)\frac{1}{\theta+2}\right) - \left((\theta+1)\frac{0}{\theta+2}\right) = \left(\frac{(\theta+1)}{\theta+2}\right) - \left((\theta+1) * 0\right) =$$

$$= \left(\frac{(\theta+1)}{\theta+2}\right) - (0) = \frac{\theta+1}{\theta+2}$$

The population first moment: $E(x) = \frac{\theta+1}{\theta+2}$

The sample first moment:

$$\frac{1}{n}\sum_{i=1}^{n} x_i^1 = \bar{x}$$

$$\frac{\theta+1}{\theta+2} = \bar{x} => \theta+1 = \theta\bar{x} + 2\bar{x} => \theta - \theta\bar{x} = 2\bar{x} - 1 => \theta(1-\bar{x}) = 2\bar{x} - 1 =>$$

$$\theta = \frac{2\bar{x} - 1}{1 - \bar{x}}$$

That is the estimator for θ

Now to compute the estimate for the sample, we first need the sample mean: $\bar{x}$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(0.92 + 0.79 + 0.9 + 0.65 + 0.86 + 0.47 + 0.73 + 0.97 + 0.84 + 0.67)$$

$$\bar{x} = \frac{1}{10} * 7.8 = 0.78$$

$$\theta = \frac{2\bar{x} - 1}{1 - \bar{x}} = \frac{2 * 0.78 - 1}{1 - 0.78} = \frac{0.56}{0.22} = 2.5454$$

$$\theta = 2.5454$$

b) Maximum Likelihood

$$L(x_1, x_2, \dots, x_{10}|\theta) = \prod_{i=1}^{10} (\theta + 1){x_i}^{\theta} = (\theta + 1)^{10} \prod_{i=1}^{10} {x_i}^{\theta}$$

Since it is hard to operate with this equation, its better to use the log-likelihood

$$\ln L(x_1, x_2, \dots, x_{10}|\theta) = l(x_1, x_2, \dots, x_{10}|\theta)$$

$$\ln(\theta + 1)^{10} \prod_{i=1}^{10} {x_i}^{\theta} = \ln(\theta + 1)^{10} + \ln \prod_{i=1}^{10} {x_i}^{\theta} = 10\ln(\theta + 1) + \sum_{i=1}^{10} \ln {x_i}^{\theta} =$$

$$10\ln(\theta + 1) + \sum_{i=1}^{10} \theta \ln x_i = 10\ln(\theta + 1) + \theta \sum_{i=1}^{10} \ln x_i$$

In order to find the maximum, we need to do the derivative and solve when it is 0.

$$l'(\theta) = \left(10\ln(\theta + 1) + \theta \sum_{i=1}^{10} \ln x_i\right) \frac{d}{d\theta}$$

$$\left(10\ln(\theta + 1) + \theta \sum_{i=1}^{10} \ln x_i\right) \frac{d}{d\theta} = [10\ln(\theta + 1)] \frac{d}{d\theta} + \left[\theta \sum_{i=1}^{10} \ln x_i\right] \frac{d}{d\theta}$$

$$[10\ln(\theta + 1)] \frac{d}{d\theta} = 10 * [\ln(\theta + 1)] \frac{d}{d\theta} = 10 * \frac{1}{\theta + 1} [\theta + 1] \frac{d}{d\theta} = \frac{10}{\theta + 1}(1 + 0) = \frac{10}{\theta + 1}$$

$$\left[\theta \sum_{i=1}^{10} \ln x_i\right] \frac{d}{d\theta} = \sum_{i=1}^{10} \ln x_i \, [\theta] \frac{d}{d\theta} = \sum_{i=1}^{10} \ln x_i * (1) = \sum_{i=1}^{10} \ln x_i$$

$$l'^{(\theta)} = \frac{10}{\theta + 1} + \sum_{i=1}^{10} \ln x_i = 0$$

$$\frac{10}{\theta + 1} + \sum_{i=1}^{10} \ln x_i = 0 \Rightarrow \frac{10}{\theta + 1} = -\sum_{i=1}^{10} \ln x_i \Rightarrow 10 = -\theta \sum_{i=1}^{10} (\ln x_i) - \sum_{i=1}^{10} (\ln x_i)$$

$$10 + \sum_{i=1}^{10} (\ln x_i) = -\theta \sum_{i=1}^{10} (\ln x_i) \Rightarrow \frac{10 + \sum_{i=1}^{10} (\ln x_i)}{\sum_{i=1}^{10} (\ln x_i)} = -\theta \Rightarrow$$

$$\theta = -\left(\frac{10 + \sum_{i=1}^{10}(\ln x_i)}{\sum_{i=1}^{10}(\ln x_i)}\right)$$

This is the MLE

Now to compute the estimate for the sample:

$$\sum_{i=1}^{10}(\ln x_i) = \ln 0.92 + \ln 0.79 + \ln 0.9 + \ln 0.65 + \ln 0.86 + \ln 0.47 + \ln 0.73 + \ln 0.97 + \ln 0.84 + \ln 0.67$$

$$\sum_{i=1}^{10}(\ln x_i) = -2.681093754$$

$$\theta = -\left(\frac{10 + \sum_{i=1}^{10}(\ln x_i)}{\sum_{i=1}^{10}(\ln x_i)}\right) = -\left(\frac{10 - 2.681093754}{-2.681093754}\right) = -\frac{7.318906246}{-2.681093754} = 2.7298$$

$$\theta = 2.7298$$

Problem 4.

$$L(x|\theta) = \frac{1}{\theta - 0} \; for\; 0 \leq x \leq \theta => \frac{1}{\theta} \; for \; 0 \leq x \leq \theta$$

When $\theta < max\{x_1, x_2, \dots, x_{10}\}$:

Since some value x_i is grater than θ that means that his $L(x_i|\theta) = 0$. And since the Likelihood function is a multiplication of its elements:

$$L(x_1, x_2, \dots, x_{10}|\theta) = \prod_{i=1}^{10} L(x_i|\theta)$$

If at least one element is 0, the result is 0, and since $\theta < max\{x_1, x_2, \dots, x_{10}\}$, that means that at least 1 element is greater than θ.

$$L(x_1, x_2, \dots, x_{10}|\theta) = 0 \; if \theta < max\{x_1, x_2, \dots, x_{10}\}$$

When $\theta \geq max\{x_1, x_2, \dots, x_{10}\}$:

$$L(x_1, x_2, \dots, x_{10}|\theta) = \prod_{i=1}^{10} \frac{1}{\theta - 0} = \left(\frac{1}{\theta}\right)^{10} \; for \; 0 \leq x_i \leq \theta \; (i = 1, \dots, 10)$$

What is the MLE:

If $\theta < max\{x_1, x_2, \dots, x_{10}\}$ that means that the likelihood function is 0, which means there are no maximums.

If $\theta \geq max\{x_1, x_2, \dots, x_{10}\}$ we need to maximize the function $L(\theta) = \left(\frac{1}{\theta}\right)^{10}$ in order to maximize it, we want the θ to be as small as possible, since de function is decreasing. But we had the condition that $\theta \geq max\{x_1, x_2, \dots, x_{10}\}$ which means that the smallest possible value for θ is $max\{x_1, x_2, \dots, x_{10}\}$