Math Statistics
Monthly HW 2

Ozaner Hansha

November 10, 2020

Question 1

Consider an exponential distribution with pdf $f_X(x) = \frac{1}{\theta}e^{-x/\theta}$.

Part a: Find the method of moments estimator of θ.

Solution: Recall that the method of moments has us set the first k moments of our distribution equal to the first k sample moments, then solve for our parameters:

$$\begin{aligned}
E[X] &= \bar{M}_1 && \text{(method of moments)} \\
\theta &= \bar{M}_1 && \text{(mean of exponential RV)} \\
&= \frac{1}{n}\sum_{i=1}^{n} X_i && \text{(def. of 1st sample moment)} \\
&= \bar{X} && \text{(def. of sample mean)}
\end{aligned}$$

And so our method of moments estimator is given by:

$$\hat{\theta}_{\text{MM}} = \bar{X}$$

Part b: Check whether the estimator from part a is biased.

Solution: Recall that the mean of our distribution (exponential) is θ and that the sample mean is always an unbiased estimator of the mean. As such, our estimator is unbiased.

Part c: Find a statistic $\hat{\theta}'$ such that $P(\theta < \hat{\theta}') = .9$.

Solution: First note the following:

$$\begin{aligned}
X_i &\sim \text{Exp}(\theta) \\
\sum_{i=1}^{n} X_i &\sim \text{Gamma}(n, \theta) && \text{(sum of i.i.d. exponential RVs is gamma)} \\
\frac{2}{\theta}\sum_{i=1}^{n} X_i &\sim \text{Gamma}(n, 2) && \text{(scaling property of gamma RV)} \\
&\sim \chi^2_{2n} && \left(\chi^2_k \sim \text{Gamma}\left(\tfrac{k}{2}, 2\right)\right)
\end{aligned}$$

And so we have the following:

$$\begin{aligned}
.9 &= P\left(\frac{2}{\theta}\sum_{i=1}^{n} X_i > \underbrace{\text{Inv-}\chi^2_{2n}(1-.9)}_{\text{cdf}}\right) && \left(\tfrac{2}{\theta}\textstyle\sum_{i=1}^{n} X_i \sim \chi^2_{2n}\right) \\
&= P\left(\frac{1}{\theta} > \frac{\text{Inv-}\chi^2_{2n}(.1)}{2\sum_{i=1}^{n} X_i}\right) \\
&= P\left(\theta < \frac{\text{Inv-}\chi^2_{2n}(.1)}{2\sum_{i=1}^{n} X_i}\right) && \text{(both sides of inequality are positive)}
\end{aligned}$$

And so our statistic $\hat{\theta}'$ is given by:

$$\hat{\theta}' = \frac{\text{Inv-}\chi^2_{2n}(.1)}{2\sum_{i=1}^n X_i} = \frac{\text{Inv-}\chi^2_{2n}(.1)}{2n\bar{X}}$$

Part d: If we measure a sample mean of $\bar{x} = 4$ with $n = 10$, produce a 90% one-sided confidence interval for the statistic in part d of the form $\theta < \hat{\theta}'(X)$.

Solution: Our measured 90% one-sided confidence interval is given by:

$$\begin{aligned}
(-\infty, \hat{\theta}'(x)) &= \left(-\infty, \frac{\chi^2_{2n}(.1)}{2n\bar{x}}\right) && \text{(part c)} \\
&= \left(-\infty, \frac{\chi^2_{20}(.1)}{2 \cdot 10 \cdot 4}\right) \\
&\approx \left(-\infty, \frac{12.44261}{80}\right) \\
&\approx (-\infty, 0.155533)
\end{aligned}$$

Question 2

Consider two distributions with the same mean μ but with different variances σ_1^2, σ_2^2. Suppose we take independent samples from each distribution of sizes n_1, n_2, and that these samples have sample means $\bar{X}_1$, $\bar{X}_2$.

Part a: Show that for any $\omega \in \mathbb{R}$, the statistic $\hat{\theta}_\omega = \omega\bar{X}_1 + (1-\omega)\bar{X}_2$ is an unbiased estimator of μ.

Solution: Let us compute the expected value of $\hat{\theta} - \omega$:

$$\begin{aligned}
E[\hat{\theta}_\omega] &= E[\omega\bar{X}_1 + (1-\omega)\bar{X}_2] && \text{(def. of } \hat{\theta}_\omega\text{)} \\
&= \omega E[\bar{X}_1] + (1-\omega)E[\bar{X}_2] && \text{(linearity of expectation)} \\
&= \omega\mu + (1-\omega)\mu && \text{(mean of sample mean)} \\
&= \omega\mu + \mu - \omega\mu \\
&= \mu
\end{aligned}$$

And so we have that, for any real ω, the mean of our estimator $\hat{\theta}_\omega$ is the mean of the population. Thus our estimator is unbiased.

Part b: Give the variance of the estimator $\hat{\theta}_\omega$.

Solution: The variance of $\hat{\theta}_\omega$ is given by:

$$\begin{aligned}
\text{Var}(\hat{\theta}_\omega) &= \text{Var}(\omega\bar{X}_1 + (1-\omega)\bar{X}_2) && \text{(def. of } \hat{\theta}_\omega\text{)} \\
&= \text{Var}(\omega\bar{X}_1) + \text{Var}((1-\omega)\bar{X}_2) && \text{(variance of independent RVs)} \\
&= \omega^2\,\text{Var}(\bar{X}_1) + (1-\omega)^2\,\text{Var}(\bar{X}_2) && \text{(variance of multiple of RV)} \\
&= \omega^2\frac{\sigma_1^2}{n_1} + (1-\omega)^2\frac{\sigma_2^2}{n_2} && \text{(variance of sample mean)}
\end{aligned}$$

Part c: Show that $\mathrm{Var}(\hat{\theta}_\omega)$ is minimized when:

$$\omega = \frac{n_1\sigma_2^2}{n_2\sigma_1^2 + n_1\sigma_2^2}$$

Solution: Our goal is to compute the following:

$$\begin{aligned}
\arg\min_{\omega\in\mathbb{R}} \mathrm{Var}(\hat{\theta}_\omega) &= \arg\min_{\omega\in\mathbb{R}} \omega^2\frac{\sigma_1^2}{n_1} + (1-\omega)^2\frac{\sigma_2^2}{n_2} && \text{(part b)} \\
&= \arg\min_{\omega\in\mathbb{R}} \underbrace{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}_{a}\omega^2 + \underbrace{\left(-\frac{2\sigma_2^2}{n_2}\right)}_{b}\omega + \underbrace{\frac{\sigma^2}{n_2}}_{c}
\end{aligned}$$

Note that the expression we are trying to minimize is a quadratic polynomial in ω. Recall that all quadratics in one variable have a single extremum, and that extremum is a minimum if $a > 0$ and a maximum if $a < 0$.

In our case, $a = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$ is always positive, unless both σ_1^2 and σ_2^2 were zero (i.e. both RVs the distributions correspond to are just constants). And so the ω that minimizes estimator's variance is given by the zero of the above quadratic:

$$\begin{aligned}
\arg\min_{\omega\in\mathbb{R}} \mathrm{Var}(\hat{\theta}_\omega) &= \arg\min_{\omega\in\mathbb{R}} \underbrace{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}_{a}\omega^2 + \underbrace{\left(-\frac{2\sigma_2^2}{n_2}\right)}_{b}\omega + \underbrace{\frac{\sigma^2}{n_2}}_{c} \\
&= -\frac{b}{2a} && \text{(zero of a quadratic)} \\
&= -\frac{-\frac{2\sigma_2^2}{n_2}}{2\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)} \\
&= \frac{\sigma_2^2}{n_2\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)} \\
&= \frac{\sigma_2^2}{\frac{n_2\sigma_1^2}{n_1} + \sigma_2^2} \\
&= \frac{n_1\sigma_2^2}{n_2\sigma_1^2 + n_1\sigma_2^2}
\end{aligned}$$

And so we are done.

Question 3

Problem: Consider an estimator $\hat{\theta}$, with mean μ and variance σ, for the parameter θ. Is the following true:

$$P(\mu - 2\sigma < \theta < \mu + 2\sigma) \approx .95$$

Solution: This is not true. The only reasonable way to interpret the given statement is by considering the parameter θ to be a constant random variable. As such, the probability given above is either 1 or 0 depending on whether $\mu - 2\sigma < \theta < \mu + 2\sigma$ is true or not. In either case, it is *not* approximately .95.

However, if $\hat{\theta} \sim \mathcal{N}(\mu, \sigma^2)$, then the following modified statement *is* true:

$$\begin{aligned}
P(\mu - 2\sigma < \hat{\theta} < \mu + 2\sigma) &= P\left(-2 < \frac{\hat{\theta} - \mu}{\sigma} < 2\right) \\
&= \Phi(2) - \Phi(-2) \approx .95 && \left(\tfrac{\hat{\theta}-\mu}{\sigma} \text{ is a standard normal RV}\right)
\end{aligned}$$

Question 4

Consider a statistic Y dependent on some paramater θ.

Part a: Suppose Y has a cdf given by:

$$F(y) = \begin{cases} 1 - \frac{1}{(\theta+y)^2}, & y > -\theta \\ 0, & \text{otherwise} \end{cases}$$

Show that $\theta \leq 10 - Y$ is a 99% confidence interval for θ.

Solution: Note the following:

$$\begin{aligned} P(\theta \leq 10 - Y) &= P(Y \leq 10 - \theta) \\ &= F_Y(10 - \theta) && \text{(def. of cdf)} \\ &= 1 - \frac{1}{(\theta + 10 - \theta)^2} && (10 - \theta > -\theta) \\ &= 1 - \frac{1}{100} = .99 \end{aligned}$$

And so $(-\infty, 10 - Y]$ is a 99% confidence interval for θ.

Part b: Suppose Y has a pdf given by:

$$f(y) = \begin{cases} \frac{3\theta}{(\theta y+1)^4}, & y > 0 \\ 0, & \text{otherwise} \end{cases}$$

Give $E[Y]$ and show that $2Y$ is an unbiased estimator of $\frac{1}{\theta}$.

Solution: The expected value of Y is given by:

$$\begin{aligned} E[Y] &= \int_0^\infty \frac{3\theta}{(\theta y + 1)^4} y \, \mathrm{d}y && \text{(def. of expected value)} \\ &= 3\theta \int_0^\infty \frac{y}{(\theta y + 1)^4} \, \mathrm{d}y \\ &= 3\theta \int_{u(0)}^{u(\infty)} \frac{u-1}{\theta u^4} \frac{\mathrm{d}u}{\theta} && \left({}^{u=\theta y+1 \to y=\frac{u-1}{\theta}}_{\frac{\mathrm{d}u}{\mathrm{d}y}=\theta \to \mathrm{d}y=\frac{\mathrm{d}u}{\theta}}\right) \\ &= \frac{3}{\theta} \int_1^\infty \frac{u-1}{u^4} \, \mathrm{d}u \\ &= \frac{3}{\theta} \left(\int_1^\infty \frac{1}{u^3} \, \mathrm{d}u - \int_1^\infty \frac{1}{u^4} \, \mathrm{d}u \right) \\ &= \frac{3}{\theta} \left(\left[-\frac{1}{2u^2} \right]_1^\infty - \left[-\frac{1}{3u^3} \right]_1^\infty \right) \\ &= \frac{3}{\theta} \left(\frac{1}{2} - \frac{1}{3} \right) \\ &= \frac{1}{2\theta} \end{aligned}$$

Now we will show that $2Y$ is an unbiased estimator of $\frac{1}{\theta}$:

$$\begin{aligned} E[2Y] &= 2E[Y] && \text{(linearity of expectation)} \\ &= 2 \cdot \frac{1}{2\theta} && \text{(see above)} \\ &= \frac{1}{\theta} \end{aligned}$$

And so it is unbiased.