MTH375: Mathematical Statistics - Homework #7

Cason Konzer

March 17, 2022

Key Concepts: Pivotal quantities, confidence intervals.

1. Let $X_1, \ldots X_{10}$ be a sample of $iid \sim Normal(0, \theta)$ random variables.

(a) Show that $T = \frac{1}{\theta}\sum_{i=1}^{n} X_i^2$ is a pivotal quantity and find its distribution.

Solution:

A pivotal quantities distribution is does not depend on θ.

- $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2_{n-1}$.
- $f_T(t) = \frac{1}{\theta}\sum_{i=1}^{10}(X_i - 0)^2 \sim \chi^2_9$.

As the chi squared distribution depends only on the sample size n, T is a pivotal quantity.

(b) Determine a 99% confidence interval for θ based on T.

Solution:

- $\mathbb{P}\Big(a \leq \frac{1}{\theta}\sum_{i=1}^{10} X_i^2 \leq b\Big) = 0.99$.
- $\mathbb{P}\Big(\frac{1}{b}\sum_{i=1}^{10} X_i^2 \leq \theta \leq \frac{1}{a}\sum_{i=1}^{10} X_i^2\Big) = 0.99$.
- `qchisq(p = c(0.005, 0.995), df = 9)` $= (1.735, 23.589) = (a, b)$.
- $L(\overrightarrow{X}) = 0.042\sum_{i=1}^{10} X_i^2$.
- $U(\overrightarrow{X}) = 0.576\sum_{i=1}^{10} X_i^2$.

(c) Suppose that $\sum_{i=1}^{10} X_i^2 = 27.3$. What is the 99% confidence interval for θ?

Solution:

- $L(\overrightarrow{X}) = 0.042 \cdot 27.3 = 1.157$.
- $U(\overrightarrow{X}) = 0.576 \cdot 27.3 = 15.735$.

Our 99% confidence interval is thus $(1.157, 15.735)$.

(d) Determine a 99% confidence interval for $\mathrm{SD}(X) = \sqrt{\theta}$.

Solution:

As we know the interval for the variance, the requested is the roots of the above bounds.

- $L^*(\overrightarrow{X}) = \sqrt{L(\overrightarrow{X})} = \sqrt{1.157} = 1.076$.
- $U^*(\overrightarrow{X}) = \sqrt{U(\overrightarrow{X})} = \sqrt{15.735} = 3.967$.

Our confidence interval is thus $(1.076, 3.967)$.

2. Let $X_1, \dots X_6$ be a sample of $iid \sim Uniform[\theta, 1]$ random variables.

(a) Show that $T = \dfrac{1 - X_{(1)}}{1 - \theta}$ is a pivotal quantity and find its cdf.

Solution:

A pivotal quantities distribution is does not depend on θ.

- $\mathbb{P}\Big(X_{(1)} > m\Big) = \Big[1 - F_{X_i}\Big]^n = \Big[1 - \dfrac{m - \theta}{1 - \theta}\Big]^n.$
- $\mathbb{P}\Big(\dfrac{1 - X_{(1)}}{1 - \theta} \leq t\Big) = \mathbb{P}\Big(-X_{(1)} \leq t(1-\theta) - 1\Big) = \mathbb{P}\Big(X_{(1)} > 1 - t(1-\theta)\Big) = \Big[1 - \dfrac{t - \theta}{1 - \theta}\Big]^n.$
- Make the substitution $t = \dfrac{1 + \theta}{t} + \theta$
- $\mathbb{P}\Big(X_{(1)} > 1 - \Big(\dfrac{1+\theta}{t} - \theta\Big)(1 - \theta)\Big) = \left[1 - \dfrac{\dfrac{1+\theta}{t} + \theta - \theta}{1 - \theta}\right]^n.$
- $\mathbb{P}\Big(X_{(1)} > 1 - \dfrac{1+\theta}{t}\Big) = \Big[1 - \dfrac{1}{t}\Big]^n.$

We can see that T is independent of θ.

(b) Determine a 85% confidence interval for θ based on T.

Solution:

- $\mathbb{P}\Big(\dfrac{1-X_{(1)}}{1-\theta} \leq a\Big) = 0.075 = a^6 \Rightarrow a = 0.075^{1/6} = 0.649.$

- $\mathbb{P}\Big(\dfrac{1-X_{(1)}}{1-\theta} \leq b\Big) = 0.925 = b^6 \Rightarrow b = 0.925^{1/6} = 0.987.$

- $\mathbb{P}\Big(0.649 \leq \dfrac{1-X_{(1)}}{1-\theta} \leq 0.987\Big) = 0.85.$

- $\mathbb{P}\Big(1 - \dfrac{1-X_{(1)}}{0.649} \leq \theta \leq 1 - \dfrac{1-X_{(1)}}{0.987}\Big) = 0.85.$

- $L(\overrightarrow{X}) = 1 - \dfrac{1-X_{(1)}}{0.649}.$

- $U(\overrightarrow{X}) = 1 - \dfrac{1-X_{(1)}}{0.987}.$

(c) Suppose that $\overrightarrow{X} = \langle 0.527, 0.803, 0.842, 0.880, 0.474, 0.558\rangle$. Find the 85% confidence interval for θ.

Solution:

- $X_{(1)} = 0.474.$

- $L(\overrightarrow{X}) = 1 - \dfrac{1-0.474}{0.649} = 0.190.$

- $U(\overrightarrow{X}) = 1 - \dfrac{1-0.474}{0.987} = 0.467.$

Our 85% confidence interval is thus $(0.190, 0.467)$.

3. Let $T = X_1, \ldots X_8$ be a sample of $iid \sim Gamma(4, \theta)$ random variables.

(a) Show that $\frac{2}{\theta}\sum_{i=1}^{n} X_i$ is a pivotal quantity and find its distribution.

Solution:

A pivotal quantities distribution is does not depend on θ.

- Let $L = \sum_{i=1}^{8} X_i \sim Gamma(32, \theta)$
- $F_L(l) = P\Big(\frac{2L}{\theta} \leq l\Big) = F_L(l) = P\Big(L \leq \frac{\theta l}{2}\Big) \sim \chi^2_{64}$.

As the chi squared distribution depends only on the sample size n, L is a pivotal quantity.

(b) Determine a 95% confidence interval for θ based on T.

Solution:

- $\mathbb{P}\Big(a \leq \sum_{i=1}^{8} X_i \leq b\Big) = 0.95$.
- $\mathbb{P}\Big(\frac{1}{b}\sum_{i=1}^{8} X_i \leq \theta \leq \frac{1}{a}\sum_{i=1}^{8} X_i\Big) = 0.95$.
- `qchisq(p = c(0.025, 0.975), df = 64)` $= (43.776, 88.004) = (a, b)$.
- $L(\overrightarrow{X}) = 0.0114\sum_{i=1}^{8} X_i$.
- $U(\overrightarrow{X}) = 0.0228\sum_{i=1}^{8} X_i$.

(c) Suppose that $\overrightarrow{X} = \langle 17.40, 15.95, 15.17, 7.92, 11.54, 10.46, 8.47, 15.54 \rangle$. Find a 95% confidence interval for θ.

Solution:

- $\displaystyle\sum_{i=1}^{8} X_i = 102.45$
- $L(\overrightarrow{X}) = 0.0114 \cdot 102.45 = 1.164.$
- $U(\overrightarrow{X}) = 0.0228 \cdot 102.45 = 2.340.$

Our 95% confidence interval is thus $(1.164, 2.340)$.

4. Let $X_1, \dots X_6$ be a sample of $iid \sim Normal(\mu, \theta)$ random variables.

(a) Why is $T = \dfrac{(n-1)S^2}{\theta}$ a pivotal quantity? What is the distribution of $\dfrac{(n-1)S^2}{\theta}$?

Solution:

T is a pivotal quantity as its distribution does not depend on θ.

- $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2_{n-1}$.
- $f_T(t) = \dfrac{1}{\theta}\sum_{i=1}^{6}(X_i - \overline{X})^2 \sim \chi^2_5$.

Conclusion.

(b) Determine a 93% confidence interval for θ based on T.

Solution:

Intro.

- $\mathbb{P}\Big(a \leq \dfrac{5S^2}{\theta} \leq b\Big) = 0.93$.
- $\mathbb{P}\Big(\dfrac{5S^2}{b} \leq \theta \leq \dfrac{5S^2}{a}\Big) = 0.93$.
- `qchisq(p = c(0.035, 0.965), df = 5)` $= (0.969, 11.985) = (a, b)$.
- $L(\overrightarrow{X}) = \dfrac{5S^2}{11.985}$.
- $U(\overrightarrow{X}) = \dfrac{5S^2}{0.969}$.

Conclusion.

(c) Suppose that $\overrightarrow{X} = \langle 0.71, 0.62, 2.28, 1.01, 10.01, 7.64 \rangle$. Find a 95% confidence interval for θ.

Solution:

- $S(\overrightarrow{X}) = 4.075$
- $S^2(\overrightarrow{X}) = 16.604$
- $L(\overrightarrow{X}) = \dfrac{5(16.604)}{11.985} = 6.927.$
- $U(\overrightarrow{X}) = \dfrac{5(16.604)}{0.969} = 85.643.$

Our 93% confidence interval is thus $(6.927, 85.643)$.

5. A pollster chooses 500 citizens at random, and asks whether each plans to vote 'yes' or 'no' on proposition X. In the sample, 375 indicate a 'yes' vote.

Find a 95% confidence interval for the proportion of the total population who plan to vote 'yes.'

Solution:

We will leveralge the general form of symmetric confidence intervals ...

$CI = sample\ statistic \pm (critical\ value)(standard\ error)$.

Given our sample is sufficently large our sampling distribution is approximately normally distributed, by the central limit theorem.

From our sample we can construct an estimate for the probablity of a voter panning to vote 'yes,' our sample statistic ...

$\hat{p} = 375/500 = 0.75$.

For the normal distribution we can solve for values (a, b), where $a = -b$ and $|a| = |b|$ is our critical value ...

`qnorm(p = c(0.025, 0.975), mean = 0, sd = 1)` $= (-1.96, 1.96) = (a, b)$.

From our sample statistic we can construct our standard error ...

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{(0.75)(0.25)}{500}} = 0.194$$

We now have the necessacary information to construct our 95% confidence interval ...

CI: $0.75 \pm (1.96)(0.194) = 0.75 \pm 0.038$.

Our 95% confidence interval is thus $(0.712, 0.788)$.