Tutorial 8

1. Let X be a single observation from a distribution with density

$$p_X(x) = \theta x^{\theta-1} \qquad 0 \leq x \leq 1.$$

   (a) Let $Y = -\frac{1}{\log X}$. Evaluate the confidence level of the interval estimator $\left[\frac{Y}{2}, Y\right]$.

   (b) Find a pivotal quantity and use it to set up an interval estimator.

2. Let $X_1, \ldots, X_n$ be a sample from a $N(\mu, \sigma^2)$ population, both μ and σ^2 unknown.

   (a) Let $Z_1, \ldots, Z_n$ be i.i.d. $N(0,1)$ variables. The distribution of $W = Z_1^2 + \ldots + Z_n^2$ is χ_n^2. Let $\underline{Z} = (Z_1, \ldots, Z_n)^t$ where t denotes transpose. Let $Y_j = Z_j - \overline{Z}$ and let $\underline{Y} = (Y_1, \ldots, Y_n)^t$. Let $\underline{Y} = M_n \underline{Z}$ for a symmetric matrix M_n. What is M_n? Prove that $M_n^2 = M_n$. From this, what do you conclude about the eigenvalues of M_n? Use the fact that the sum of the eigenvalues is equal to the sum of the trace for symmetric matrices.

   Now consider the expression $M_n = PDP^t$ where D is diagonal and P is orthonormal. What is the distribution of $P^t \underline{Z}$? Hence what is the distribution of $\underline{Z}^t M_n^t M_n \underline{Z}$?

   Hence conclude that

$$\frac{\sum_{j=1}(X_j - \overline{X})^2}{\sigma^2} \sim \chi_{n-1}^2.$$

   (b) Show that

$$\overline{X} \perp X_1 - \overline{X}, \ldots, X_n - \overline{X}.$$

3. Let $Z \sim N(0,1)$ and let $V \sim \chi_m^2$. Let $Z \perp V$. Let $T = \frac{Z}{\sqrt{V/m}}$. Compute the density function for T.

4. Let $X_1, \ldots, X_n$ be a random sample from a $N(1, \sigma^2)$ population ($\mu = 1$ is known). Construct a symmetric $1 - \alpha$ interval estimator for σ, with as many degrees of freedom as possible.

5. Let $X_1, \ldots, X_n$ be independent $N(\mu, \sigma^2)$ random variables. Using a pivot based on $\sum_{j=1}^n (X_j - \overline{X})^2$, construct a symmetric confidence interval with confidence level $1 - \alpha$ for $\log \sigma^2$.

6. Suppose that $Y_1, \ldots, Y_n$ are independent and that

$$Y_i \sim N(x_i \beta, \sigma^2),$$

   where $x_1, \ldots, x_n$ are given, σ is known and β is an unknown parameter.

   (a) Compute the least squares estimator $\widehat{\beta}_{LS}$ of β.

   (b) Compute a confidence interval for β of the form $[\widehat{\beta}_{LS} - c, \widehat{\beta}_{LS} + c]$ with confidence level $1 - \alpha = 0.95$.

7. Let

$$X_i = \frac{\theta}{2} t_i^2 + \epsilon_i \qquad i = 1, \ldots, n$$

where ϵ_i are i.i.d. $N(0, \sigma^2)$ variables, where σ is known.

(a) Compute the MLE of θ.

(b) Using a pivot based on the MLE of θ, find a symmetric confidence interval for θ with confidence level $1 - \alpha$.

(c) Suppose the values for t_i may be chosen freely subject to the constraint that $0 \leq t_i \leq 1$ for each $i = 1, \ldots, n$. What values of t_i should be chosen to make the symmetric confidence interval as short as possible?

8. Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma_1^2)$. Suppose a lower confidence bound $\overline{X} - c$, intended to be of confidence level $1 - \alpha$ is computed under the assumption that $X_j \sim N(\mu, \sigma_0^2)$. What is the actual confidence level?

9. (a) Let $X_1, \ldots, X_n$ be a $N(\mu, \sigma^2)$ sample, where μ and σ^2 are both unknown. Show that the symmetric $1 - \alpha$ confidence interval is given by

$$\left[\overline{X} \pm \frac{S}{\sqrt{n}} t_{n-1;\alpha/2} \right]$$

where $S^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - \overline{X})^2$ and $t_{n,\alpha}$ denotes the number such that $\mathbb{P}(T > t_{n,\alpha}) = \alpha$ where $T \sim t_n$.

(b) Suppose we want to select a sample size N such that the interval in part (a) has length at most $l = 2d$ for some preassigned length. Stein's two stage procedure (1945) is the following: Begin by taking a fixed number $n_0 \geq 2$ of observations, calculate $\overline{X}_0 = \frac{1}{n_0} \sum_{j=1}^{n_0} X_j$ and $S_0^2 = \frac{1}{n_0 - 1} \sum_{j=1}^{n_0} (X_j - \overline{X}_0)^2$. Then take $N - n_0$ further observations where N is the smallest integer greater than or equal to n_0 and greater than or equal to $(S_0 t_{n_0-1;(\alpha/2)}/d)^2$.

Show that

$$\frac{\sqrt{N}(\overline{X} - \mu)}{S_0} \sim t_{n_0-1}$$

where $S_0^2 = \frac{1}{n_0-1} \sum_{j=1}^{n_0} (X_j - \overline{X}_0)^2$ and $\overline{X} = \frac{1}{N} \sum_{j=1}^{N} X_j$. It follows that $\left[\overline{X} \pm \frac{S_0}{\sqrt{N}} t_{n_0-1;\alpha/2} \right]$ is a confidence interval with confidence level $1 - \alpha$ for μ of length at most $2d$.

Hint recall that $\overline{X} \perp S^2$ where $\overline{X}$ is the estimator of μ and S^2 is the estimator of σ^2 based on a sample size n from a $N(\mu, \sigma^2)$ distribution. Consider the definition of a t distribution.

10. Let $X_1, \ldots, X_n$ be a random sample from a Rayleigh distribution

$$p(x, \sigma) = \begin{cases} \frac{x}{\sigma^2} \exp\left\{ -\frac{x^2}{2\sigma^2} \right\} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\sigma > 0$ is an unknown parameter.

(a) Compute the maximum likelihood estimator of σ^2.

(b) Compute a symmetric confidence interval, confidence level $1-\alpha$ of the form

$$\left[c_n\widehat{\sigma^2}_{ML}, d_n\widehat{\sigma^2}_{ML}\right]$$

for σ^2. Express your answer in terms of quantiles of an appropriate χ^2 distribution.

11. Let $X_1,\ldots,X_n$ be a $N(\mu,\sigma^2)$ random sample where σ is known. Show that the interval of shortest length of confidence $1-\alpha$ of the form

$$\left[\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha_1}, \overline{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha_2}\right] \qquad \alpha_1+\alpha_2=\alpha$$

is obtained by taking $\alpha_1=\alpha_2=\frac{\alpha}{2}$.