Data Science in Aerospace

Pontual and Interval Estimation

Latest Update: *16th May, 2025, 19:10*

ISEC Lisboa — Escola de Gestão, Engenharia e Aeronáutica

1 Aircraft Altitude

Consider the altitude H of an aircraft with a probability density function given by

$$f(h) = \begin{cases} |h| & \text{if } |h| < 1, \\ 0 & \text{if } |h| \geq 1, \end{cases}$$

and a simple random sample of altitude measurements $(H_1, H_2, H_3, H_4, H_5)$.

- Determine the probability density function of the simple random sample.

$$\begin{aligned} f_{\mathbf{h}}(h_1, h_2, h_3, h_4, h_5) &= f_{h_1}(h_1) f_{h_2}(h_2) f_{h_3}(h_3) f_{h_4}(h_4) f_{h_5}(h_5) \\ &= f(h_1) f(h_2) f(h_3) f(h_4) f(h_5) \\ &= \begin{cases} |h_1||h_3||h_3||h_4||h_5| & \text{if } |h_i| \leq 1, \forall i \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

- Determine the expected value and the variance of the mean altitude from the simple random sample.

$$\overline{H} = \frac{1}{5}(H_1 + H_2 + H_3 + H_4 + H_5)$$

$$E[\overline{H}] = \frac{1}{5}(E[H_1] + E[H_2] + E[H_3] + E[H_4] + E[H_5]) = E[H] = 0$$

$$E[H] = \int_{-1}^{0} -h^2 \, dh + \int_{0}^{1} h^2 \, dh = 0$$

$$
\begin{aligned}
\mathrm{Var}(\overline{H}) &= \frac{1}{5^2}(\mathrm{Var}(H_1) + \mathrm{Var}(H_2) + \mathrm{Var}(H_3) + \mathrm{Var}(H_4) + \mathrm{Var}(H_5)) \\
&= \frac{1}{5^2} \times 5\mathrm{Var}(H) \\
&= \frac{1}{5}\mathrm{Var}(H) = \frac{1}{10}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(H) &= \int_{-1}^{0} (h-0)^2(-h)\ dh + \int_{0}^{1} (h-0)^2 h\ dh \\
&= \int_{-1}^{0} -h^3\ dh + \int_{0}^{1} h^3\ dh \\
&= \frac{1}{2}
\end{aligned}
$$

2 Aircraft Performance Estimators

Let $\hat{\Theta}_1$ and $\hat{\Theta}_2$ be estimators of the true fuel efficiency θ (measured in miles per gallon) of an aircraft, such that:

$$
E(\hat{\Theta}_1) = \theta, \quad \mathrm{Var}(\hat{\Theta}_1) = 9
$$
$$
E(\hat{\Theta}_2) = 3\theta, \quad \mathrm{Var}(\hat{\Theta}_2) = 3
$$

- Determine and justify which of these estimators is the better estimator for θ.

$$
\mathrm{MSE} = \left(E[\hat{\Theta}] - \theta\right)^2 + \mathrm{Var}(\hat{\Theta})
$$

$$
\mathrm{MSE}_1 = (\theta - \theta)^2 + \mathrm{Var}(\hat{\Theta}_1) = 9
$$

$$
\mathrm{MSE}_2 = (3\theta - \theta)^2 + \mathrm{Var}(\hat{\Theta}_2) = 4\theta^2 + 3
$$

$$
\begin{aligned}
&\mathrm{MSE}_2 > \mathrm{MSE}_1 \\
&4\theta^2 + 3 > 9 \\
&\qquad \theta > \sqrt{3/2} \text{ or } -\theta > \sqrt{3/2} \\
&\qquad \theta > \sqrt{3/2} \text{ or } \theta < -\sqrt{3/2}
\end{aligned}
$$

The estimator $\hat{\Theta}_2$ is more eficient when $|\theta| < \sqrt{3/2}$

3 Maximum Cruise Speed

An aircraft's cruising speed V is uniformly distributed in the interval $[0, \theta]$. The goal is to estimate the maximum possible cruising speed θ based on random samples of observed speeds $\tilde{V}$ from n flights. Consider the following estimators for θ:

1. $\hat{\Theta}_1$, which is the maximum observed cruising speed.
2. $\hat{\Theta}_2 = 2\bar{V}$, where $\bar{V}$ is the sample mean of the observed speeds.

- Determine, with justification, which of these estimators is a better estimator for θ knowing that

$$E(\hat{\Theta}_1) = \frac{n}{n+1}\theta, \quad \text{Var}(\hat{\Theta}_1) = \frac{n}{(n+1)^2}\theta^2$$

$$E[V] = \frac{\theta}{2}$$

$$\text{Var}(V) = \theta^2/12$$

$$E[\hat{\Theta}_2] = E[2\overline{V}] = 2E[V] = \theta$$

$$\text{Var}(\hat{\Theta}_2) = \text{Var}(2\overline{V}) = 4\text{Var}(\overline{V}) = 4\frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

$$\begin{aligned}
\text{MSE}_2 &> \text{MSE}_1 \\
\left(E[\hat{\Theta}_2] - \theta\right)^2 + \text{Var}(\hat{\Theta}_2) &> \left(E[\hat{\Theta}_1] - \theta\right)^2 + \text{Var}(\hat{\Theta}_1) \\
(\theta - \theta)^2 + \frac{\theta^2}{3n} &> \left(\theta\frac{n}{n+1} - \theta\right)^2 + \theta^2\frac{n}{(n+1)^2} \\
\frac{\theta^2}{3n} &> \theta^2\left(\frac{n}{n+1} - 1\right)^2 + \theta^2\frac{n}{(n+1)^2} \\
\frac{1}{3n} &> \left(\frac{n}{n+1} - 1\right)^2 + \frac{n}{(n+1)^2} \\
\frac{1}{3n} &> \left(\frac{n-(n+1)}{n+1}\right)^2 + \frac{n}{(n+1)^2} \\
\frac{1}{3n} &> \frac{1}{(n+1)^2} + \frac{n}{(n+1)^2} \\
\frac{1}{3n} &> \frac{1+n}{(n+1)^2} \\
\frac{1}{3n} &> \frac{1}{(n+1)} \\
3n &< n+1 \\
2n &< 1 \\
n &< 1/2
\end{aligned}$$

The estimator $\hat{\Theta}_2$ is more eficient

4 Estimating Aircraft Proportions

Consider an airport with planes painted in two colors: white and blue, in the proportion of 3:1. However, it is unknown which color is dominant. Let p represent the probability of a plane being blue during a random observation.

- What is the maximum likelihood estimate for p if, after observing 3 planes (with replacement), we find:
  - 1 blue plane?

$$X \sim B(3, p)$$

$$\mathcal{L}(1/4) = 3 \times \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{27}{64}$$

$$\mathcal{L}(3/4) = 3 \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{9}{64}$$

$p = 1/4$ is the maximum likelihood estimate

– 2 blue planes?

$$X \sim B(3, p)$$

$$\mathcal{L}(1/4) = 3 \times \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{64}$$

$$\mathcal{L}(3/4) = 3 \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{27}{64}$$

$p = 3/4$ is the maximum likelihood estimate

– Suppose now that we have no prior knowledge about the proportion of white and blue planes. What is the maximum likelihood estimate for p if, after observing 3 planes (with replacement), we find 2 blue planes?

$$X \sim B(3, p)$$

$$\mathcal{L} = 3 \times p^2(1 - p) = 3p^2 - 3p^3$$

$$\frac{d\mathcal{L}}{dp} = 6p - 9p^2$$

$$\frac{d\mathcal{L}}{dp} = 0 \Rightarrow p = 2/3$$

$p = 2/3$ is the maximum likelihood estimate

5 Quality Control

In a routine quality control process for the production of aircraft components, 4 batches of 80 components each were analyzed. The percentages of defective components in these batches were found to be 2.5%, 3.75%, 5%, and 6.25%, respectively. Assume the distribution of the number of defective components per batch follows a binomial distribution.

- Deduce the maximum likelihood estimator (MLE) for the probability that a component is defective for each batch.
- Calculate the MLE based on the sample of 4 batches.

6 Structural Tests

Suppose that the airspeed at which an aircraft's wing experiences structural failure follows a Normal distribution. For a sample of 12 aircraft wings, structural failures occurred at the following airspeeds (in knots):

$$52, \ 64, \ 38, \ 68, \ 66, \ 52, \ 60, \ 44, \ 48, \ 46, \ 70, \ 62$$

- Determine the maximum likelihood estimates of the following parameters:
  - The expected value (mean) of the airspeed.

$$V_F \sim N(\mu, \sigma^2)$$

$$f(x_i \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) \\ &= \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right). \end{aligned}$$

$$\ell(\mu, \sigma^2) = \log \mathcal{L}(\mu, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2} \cdot 2\sum_{i=1}^{n}(x_i - \mu)(-1) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \quad \implies \quad \sum_{i=1}^{n}(x_i - \mu) = 0$$

The MLE for the expected value is $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$

- The variance of the airspeed.

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

$$-n\sigma^2 + \sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

The MLE for the variance is $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$

- The standard deviation of the airspeed.

$$\sigma = \sqrt{\hat{\sigma}^2}$$

- The probability that an aircraft wing can withstand airspeeds higher than the maximum airspeed recorded in the sample above.

$$V_F \sim \mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$$

$$P(V_F \geq 70)$$

7 Aircraft Components

Measurements of the lengths of 25 components produced for an aircraft led to a sample mean of $\bar{x} = 140\,\text{mm}$. Assume that each component has a random length following a normal distribution with an expected value μ and a standard deviation $\sigma = 10\,\text{mm}$, and that the length of each component is independent of the others.

- Construct a 95% confidence interval for the expected value of the population.

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

$$\left[\overline{x} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \overline{x} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right]$$

$$\left[140 - \frac{10}{\sqrt{25}} z_{0.025}, 140 + \frac{10}{\sqrt{25}} z_{0.025}\right]$$

$$z_{0.025} \approx 1.96$$

$$\mu \in [136.1, 143.9] \text{ mm}$$

8 Air Polution

Studies were conducted at a busy airport to determine the concentration of carbon monoxide near the runways. For this purpose, air samples were collected, and their respective concentrations were determined using a spectrometer. The measurement results in ppm (parts per million) over a year were as follows:

102.2	98.4	104.1	101.0	102.2	100.4	98.60	88.20	78.80	83.00
84.70	94.8	105.1	106.2	111.2	108.3	105.2	103.2	99.00	98.80

- Determine a 95% confidence interval for the expected concentration of carbon monoxide, as well as for its variance. Indicate the hypotheses considered.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \approx 98.7 \text{ ppm}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \approx 75.5 \text{ ppm}^2$$

$$S \approx 8.7 \text{ ppm}$$

– Does the population follow a normal distribution?

$$\left[\overline{x} - \frac{S}{\sqrt{n}}t_{\alpha/2,n-1}, \overline{x} + \frac{S}{\sqrt{n}}t_{\alpha/2,n-1}\right]$$

$$\left[98.7 - \frac{8.7}{\sqrt{20}}t_{0.05/2,19}, 98.7 + \frac{8.7}{\sqrt{20}}t_{0.05/2,19}\right]$$

$$\approx [94.6, 102.8] \text{ ppm}$$

– Variance assuming that the population follows a normal distribution

$$\left[\frac{S^2(n-1)}{\chi^2_{\alpha/2,n-1}}, \frac{S^2(n-1)}{\chi^2_{1-\alpha/2,n-1}}\right]$$

$$\left[\frac{75.5 \times (20-1)}{\chi^2_{0.05/2,19}}, \frac{75.5 \times (20-1)}{\chi^2_{1-0.05/2,19}}\right]$$

$$\approx [43.4, 161.1] \text{ ppm}^2$$

9 Fuel Consumption Rate

Suppose that the fuel consumption rate of an aircraft engine at a specific power configuration, measured in liters per minute, is a random variable with a normal distribution. A sample of size 12 of this random variable yielded the following results:

2.3, 1.9, 2.1, 2.8, 2.3, 3.6, 1.4, 1.8, 2.1, 3.2, 2.0, 1.9.

- Construct a 99% confidence interval for:

– The expected fuel consumption rate of the aircraft engine.

$$\left[\overline{x} - \frac{S}{\sqrt{n}} t_{\alpha/2,n-1}, \overline{x} + \frac{S}{\sqrt{n}} t_{\alpha/2,n-1}\right]$$

$$\left[2.28 - \frac{0.60}{\sqrt{12}} t_{0.01/2,11}, 2.28 + \frac{0.60}{\sqrt{12}} t_{0.01/2,11}\right]$$

$$\approx [1.74, 2.82] \text{ l/m}$$

– The standard deviation of the fuel consumption rate of the aircraft engine.

$$\left[\sqrt{\frac{S^2(n-1)}{\chi^2_{\alpha/2,n-1}}}, \sqrt{\frac{S^2(n-1)}{\chi^2_{1-\alpha/2,n-1}}}\right]$$

$$\left[\sqrt{\frac{0.36 \times (12-1)}{\chi^2_{0.01/2,11}}}, \sqrt{\frac{0.36 \times (12-1)}{\chi^2_{1-0.01/2,11}}}\right]$$

$$\approx [0.38, 1.23] \text{ l/m}$$

10 Maintenance Procedure Analysis

In the context of a study on a specific aircraft maintenance procedure, 25 observations of the time required to complete the procedure were collected. The sample standard deviation was 0.3 hours.

- Construct the confidence intervals at 90%, 95%, and 99% for the variance of the maintenance times, indicating the underlying assumptions for constructing these intervals.

11 Structural Testing

The engineering department of an aviation company aims to estimate the breaking stress of a certain type of aircraft component. Based on a large set of tests conducted in the past, it is estimated that the standard deviation of the breaking stress of the material is 70 psi. It is desired to define a 99% confidence interval for the expected value of the breaking stress, ensuring that its width does not exceed 60 psi.

- What is the number of tests required to define this interval?

12 Aircraft Fuel Consumption

An aircraft's fuel consumption rate, X, follows a normal distribution with a mean μ (liters per hour) and a standard deviation $\sigma = 2$ liters per hour. A random sample of 25 flights was taken, and the sample mean fuel consumption rate was found to be $\bar{x} = 78.3$ liters per hour.

- Calculate the 99% confidence interval for μ.
- What is the maximum error of estimation (at 99% confidence) when estimating μ using $\bar{x} = 78.3$?
- What should the sample size be so that the maximum error of estimation, at 99% confidence, does not exceed $\epsilon = 0.1$ liters per hour?

13 Sulfur Content in Fuel

A process to determine the sulfur content in aviation fuel provided the following results (measured in percentage by weight):

$$1.12,\ 1.10,\ 1.08,\ 1.06,\ 1.08,\ 1.14,\ 1.10,\ 1.11,\ 1.14$$

- Find a 95% confidence interval for the true mean sulfur content in the aviation fuel, stating any assumptions you make about the population distribution.
- According to certain aviation standards, the sulfur content in fuel must not have a standard deviation greater than 0.02. Does the collected sample allow you to assert that the fuel complies with these standards? Justify your answer.
- Another aviation fuel supplier collected a sample of 20 measurements and reported a standard deviation of 0.085. Can you conclude that the variability of sulfur content in this supplier's fuel is three times higher than that of the first supplier?

14 Airport Delays

The following values represent the delays of flights, expressed in minutes, observed at a certain airport:

$$2.23,\ 7.74,\ 1.03,\ 0.08,\ 11.14,\ 12.72,\ 0.42,\ 5.17$$

- Assuming that the delays follow a distribution $U(0, \theta)$ and that the sample is random, calculate the maximum likelihood estimate of the parameter θ.

15 Surface Treatment Analysis

An aviation researcher wants to determine whether a specific maintenance treatment applied to aircraft surfaces affects the amount of material removed during a cleaning process. A random sample of 100 untreated aircraft surfaces was cleaned for 24 hours, resulting in an average material removal of 12.2 grams with a standard deviation of 1.1 grams. A second random sample of 200 treated aircraft surfaces was cleaned for the same duration, resulting in an average material removal of 9.1 grams with a standard deviation of 0.9 grams.

- Determine a 98% confidence interval for the difference between the true mean amounts of material removed from untreated and treated surfaces.
- Does the treatment reduce the amount of material removed?