Parametric Statistics

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

Lecture Summary

Example

Data on calorie content in 20 different beef hot dogs from Consumer Reports (June 1986 issue):

$$186, 181, 176, 149, 184, 190, 158, 139, 175, 148,$$

$$152, 111, 141, 153, 190, 157, 131, 149, 135, 132$$

- $\overline{X}_n = 156.85$, $\hat{\sigma}^2_{MLE}$
- Let's say I want to answer $P(|\overline{X}_n - \mu| < 5)$.
- If we know σ^2, use CLT.

$$Z = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \sim \mathcal{N}(0,1)$$

- If we don't know σ^2?

The t distributions

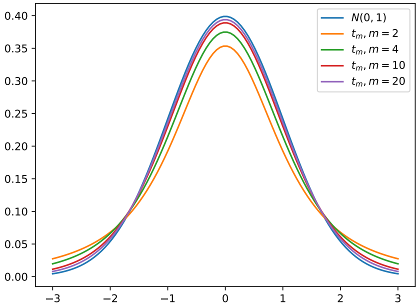
Let $Y \sim \chi^2_m$ and $Z \sim \mathcal{N}(0,1)$ be independent. Then the distribution of $X = \frac{Z}{\left(\frac{Y}{m}\right)^{1/2}}$ is called the t distribution with m degrees of freedom, or t_m.

- Pdf of the t distribution:

$$\frac{\Gamma(\frac{m+1}{2})}{(m\pi)^{1/2}\Gamma(\frac{m}{2})}(1+\frac{x^2}{m})^{-(m+1)/2}, -\infty < x < \infty$$

- No closed form CDF, tabulated at the end of statistics books

Relation to the normal distribution



- If $X \sim t_m$ then
  - $E(X) = 0$ if $m > 1$, does not exist otherwise.
  - $Var(X) = \frac{m}{m-2}$ if $m - 2 > 0$, does not exist otherwise.
  - As $n \to \infty$, t_n converges in pdf to $\mathcal{N}(0, 1)$.

Relation to samples of a normal distribution

Theorem (8.4.2)

Let $X_1, \dots, X_n$ be a random sample from $\mathcal{N}(\mu, \sigma^2)$ and let $\overline{X}_n$ be the sample mean, and define

$$S_n = \left(\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}\right)^{1/2}$$

Then $n^{1/2}(\overline{X}_n - \mu)/S_n)$ follows the t distribution with $n-1$ degrees of freedom.

- Notice that S_n is not the MLE for σ, but $\left(\frac{n-1}{n}\right)^{1/2}\hat{\sigma}_{MLE}$
- For large n, $\hat{\sigma}_{MLE}$ and S_n are close.

Review

- Let $X_1, \dots, X_n$ be a random sample from $\mathcal{N}(\mu, \sigma^2)$
- If you know μ but not σ^2

$$\frac{n\hat{\sigma}^2_{MLE}}{\sigma^2} \sim \chi^2_n, \text{ where } \hat{\sigma}^2_{MLE} \text{ is the MLE for } \sigma^2$$

- If you do not know μ or σ^2, then

$$\frac{nS_n}{\sigma^2} \sim \chi^2_{n-1}, \text{ where } S_n = \frac{\sum (X_i - \overline{X}_n)^2}{n} \text{ is the MLE for } \sigma^2$$

$$n^{1/2}(\overline{X}_n - \mu)/S_n \sim t_{n-1}, \text{ where } S_n = \left(\frac{\sum (X_i - \overline{X}_n)^2}{n-1}\right)^{1/2}$$

Back to our Example

Data on calorie content in 20 different beef hot dogs from Consumer Reports (June 1986 issue):

$$186, 181, 176, 149, 184, 190, 158, 139, 175, 148,$$

$$152, 111, 141, 153, 190, 157, 131, 149, 135, 132$$

- $\overline{X}_n = 156.85$, $S_n = 98.69$
- How confident am I in my $\hat{\mu}$ estimate?
- I know that

$$U = \frac{n^{1/2}(\overline{X}_n - \mu)}{S_n} \sim t_{n-1}$$

- I can compute $P(-c < U < c)$.

Confidence Intervals

- I can compute

$$P(\overline{X}_n - \frac{cS_n}{n^{1/2}} < \mu < \overline{X}_n + \frac{cS_n}{n^{1/2}})$$

Definition (Confidence Interval)

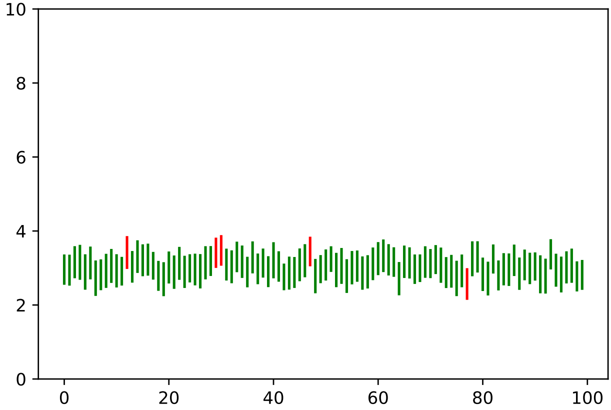
Let $X_1, \dots, X_n$ be a random sample from $f(x|\theta)$, where θ is unknown. Let $g(\theta)$ be a real-valued function, and let A and B be statistics where $P(A < g(\theta) < B) \geq \gamma \quad \forall\theta$. Then the random interval (A, B) is called a $100\gamma\%$ confidence interval for $g(\theta)$. If equality holds, the CI is exact.

- Notice: A, B are random variables.
- After a random sample is observed, A, B take specific values a and b. The interval (a, b) is then called the observed value of the confidence interval.

Confidence Intervals: Interpretation

- After observing our sample, we find that (a, b) is our 95%-CI for μ.
- This does not mean that $P(a < \mu < b) = 0.95$. In fact, we can not make such statements if we consider μ to be a number (frequentist view).
- We can think of our interpretation as repeated samples.
  - Take a random sample of size n from $\mathcal{N}(\mu, \sigma^2)$.
  - Compute (a, b).
  - Repeat many times.
  - There is a 95% chance for the random intervals to include the value of μ.

Confidence Intervals - the zipper plot



Confidence Intervals

- More generally we want to find $P(c_1 < U < c_2) = \gamma$
- Symmetric confidence intervals: Equal probability on both sides: $P(U \leq c_1) = P(U \geq c_2) = \frac{1-\gamma}{2}$
- One-sided confidence interval: All the extra probability is on one side.
- $c_1 = -\infty$ or $c_2 = \infty$.

One-sided Confidence Intervals

Definition (Lower Confidence Limit)

Let A be a statistic so that

$$P(A < g(\theta)) \geq \gamma \quad \forall\theta$$

The random interval (A, ∞) is a one-sided $100\gamma\%$ confidence interval for $g(\theta)$.
A is a $100\gamma\%$ lower confidence limit for $g(\theta)$

Definition (Upper Confidence Limit)

Let B be a statistic so that

$$P(g(\theta) < B) \geq \gamma \quad \forall\theta$$

The random interval $(-\infty, B)$ is a one-sided $100\gamma\%$ confidence interval for $g(\theta)$.
B is a $100\gamma\%$ upper confidence limit for $g(\theta)$

Pivotal

Definition (Pivotal)

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from a distribution that depends on parameter θ. Let $V(\mathbf{X}, \theta)$ be a random variable whose distribution is the same for all θ. Then V is called a pivotal quantity.

To use this we need to be able to invert the pivotal relationship: find a function $r(v, \mathbf{x})$ so that

$$r(V(\mathbf{X}, \theta), \mathbf{X}) = g(\theta)$$

If the r function is increasing in v for every $\mathbf{x}, V$ has a continuous distribution with cdf $F(v)$ and $\gamma_2 - \gamma_1 = \gamma$, then

$$\boldsymbol{A} = r\left(F^{-1}\left(\gamma_1\right), \mathbf{X}\right) \quad \text{and} \quad B = r\left(F^{-1}\left(\gamma_2\right), \mathbf{X}\right)$$

are the endpoints of an exact $100\gamma\%$ confidence interval (Theorem 8.5.3).

Pivotal Example

Variance of the normal distribution $X_1, ..., X_n$ i.i.d. $N(\mu, \sigma^2)$, both unknown.

- Find a symmetric $\gamma\%$ confidence interval for σ^2
- Find the observed symmetric $\gamma\%$ confidence interval for σ^2 for the hotdog example

Practice Exercises

8.4	2
8.5	2,3,4,7