Chapter 8: Sampling distributions of estimators

Sections

- 8.1 Sampling distribution of a statistic
- 8.2 The Chi-square distributions
- 8.3 Joint Distribution of the sample mean and sample variance
  - Skip: p. 476 - 478
- 8.4 The *t* distributions
  - Skip: derivation of the pdf, p. 483 - 484
- 8.5 Confidence intervals
- 8.6 Bayesian Analysis of Samples from a Normal Distribution
- 8.7 Unbiased Estimators
- 8.8 Fisher Information

Bayesian alternative to confidence intervals

- Bayesian inference is the posterior distribution.
- Reporting a whole distribution may not be what you (or your client) want

Bayesian alternative to confidence intervals

- Bayesian inference is the posterior distribution.
- Reporting a whole distribution may not be what you (or your client) want
- For point estimates we use Bayesian estimators
  - Minimize the expected loss

Bayesian alternative to confidence intervals

- Bayesian inference is the posterior distribution.
- Reporting a whole distribution may not be what you (or your client) want
- For point estimates we use Bayesian estimators
  - Minimize the expected loss
- If we want an interval to go with the point estimator we simply use quantiles of the posterior distribution
- For example: We can find constants c_1 and c_2 so that

$$P(c_1 < \theta < c_1 | \mathbf{X} = \mathbf{x}) \geq \gamma$$

- The interval (c_1, c_2) is called a $100\gamma\%$ *Credible interval for* θ

Bayesian alternative to confidence intervals

- Bayesian inference is the posterior distribution.
- Reporting a whole distribution may not be what you (or your client) want
- For point estimates we use Bayesian estimators
  - Minimize the expected loss
- If we want an interval to go with the point estimator we simply use quantiles of the posterior distribution
- For example: We can find constants c_1 and c_2 so that

$$P(c_1 < \theta < c_1 | \mathbf{X} = \mathbf{x}) \geq \gamma$$

- The interval (c_1, c_2) is called a $100\gamma\%$ *Credible interval for* θ
- Note: The interpretation is very different from interpretation of confidence intervals

Bayesian Analysis for the normal distribution

Let $X_1, \ldots, X_n$ be a random sample for $N(\mu, \sigma^2)$

In Chapter 7.3 we saw:

- If σ^2 is known, the normal distribution is a conjugate prior for μ
- Theorem 7.3.3: If the prior is $\mu \sim N(\mu_0, \nu_0^2)$ the posterior of μ is also normal with mean and variance

$$\mu_1 = \frac{\sigma^2 \mu_0 + n\nu_0^2 \overline{x}_n}{\sigma^2 + n\nu_0^2} \quad \text{and} \quad \nu_1^2 = \frac{\sigma^2 \nu_0^2}{\sigma^2 + n\nu_0^2}$$

- We can obtain credible intervals for μ from this $N(\mu_1, \nu_1^2)$ posterior distribution

Bayesian Analysis for the normal distribution

Let $X_1, \ldots, X_n$ be a random sample for $N(\mu, \sigma^2)$

In Chapter 7.3 we saw:

- If μ is known, the Inverse-Gamma distribution is a conjugate prior for σ^2
- Example 7.3.15: If the prior is $\sigma^2 \sim \mathrm{IG}(\alpha_0, \beta_0)$ the posterior of σ^2 is also Inverse-Gamma with parameters

$$\alpha_1 = \alpha_0 + \frac{n}{2} \quad \text{and} \quad \beta_1 = \beta + \frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2$$

- We can obtain credible intervals for σ^2 from this $\mathrm{IG}(\alpha_1, \beta_1)$ posterior distribution

What if both μ and σ^2 are unknown?

- We need the joint posterior distribution of μ and σ^2

Bayesian Analysis for the normal distribution

When both μ and σ^2 are unknown

Def: Precision

The *precision* of a normal distribution is the reciprocal of the variance:

$$\tau = \frac{1}{\sigma^2}$$

- It is somewhat simpler to work with the precision than the variance
- The pdf of the normal distribution is then written as

$$f(x|\mu, \tau) = \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\tau(x-\mu)^2\right)$$

Normal-Gamma distribution

Def: Normal-Gamma distribution

Let μ and τ be random variables where $\mu|\tau$ has the normal distribution with mean μ and precision $\lambda_0\tau$ and τ has the Gamma distribution:

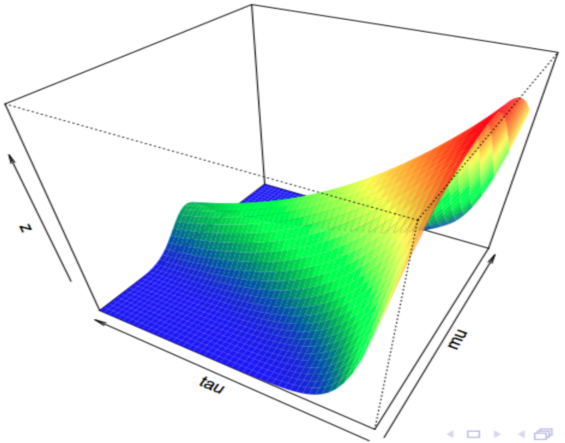
$$\mu|\tau \sim N(\mu_0, 1/\lambda_0\tau) \quad \text{and} \quad \tau \sim \mathrm{Gamma}(\alpha_0, \beta_0)$$

Then the joint distribution of μ and τ is called the *Normal-Gamma distribution with parameters μ_0, λ_0, α_0 and β_0*. The pdf for this distribution is

$$f(\mu, \tau|\mu_0, \lambda_0, \alpha_0, \beta_0) = \frac{(\lambda_0\tau)^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\lambda_0\tau(\mu-\mu_0)^2\right) \times \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\tau^{\alpha_0-1}\exp\left(-\beta_0\tau\right)$$

Pdf of the Normal-Gamma distribution

$\mu_0 = 0$, $\lambda_0 = 1$, $\alpha_0 = 0.5$ and $\beta_0 = 0.5$



Bayesian Analysis for the normal distribution

When both μ and σ^2 are unknown

Theorem 8.6.1: Conjugate prior for μ and τ

Let $X_1, \dots, X_n$ be a random sample from $N(\mu, 1/\tau)$. The Normal-Gamma distribution with parameters μ_0, λ_0, α_0 and β_0 is a conjugate prior distribution for (μ, τ) and the posterior distribution has (hyper)parameters

$$\mu_1 = \frac{\lambda_0 \mu_0 + n\overline{x}_n}{\lambda_0 + n}, \qquad \lambda_1 = \lambda_0 + n$$

$$\alpha_1 = \alpha_0 + \frac{1}{2} \qquad \text{and } \beta_1 = \beta_0 + \frac{1}{2} s_n^2 + \frac{n\lambda_0(\overline{x}_n - \mu_0)^2}{2(\lambda_0 + n)}$$

where

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad s_n^2 = \sum_{i=1}^{n}(x_i - \overline{x}_n)^2$$

Bayesian Analysis for the normal distribution

- To give credible intervals for μ and σ^2 individually we need the marginal posterior distributions
- From the structure of the Normal-Gamma distribution we immediately get the marginal for τ:

$$\tau|\mathbf{X} = \mathbf{x} \sim \text{Gamma}(\alpha_1, \beta_1)$$

This distribution can be used to obtain credible intervals for τ, or any function of τ.

Marginal distribution of the mean

Theorem 8.6.2: The marginal distribution of μ

Let the joint distribution of μ and τ be the Normal-Gamma distribution with parameters μ_0, λ_0, α_0 and β_0. Then

$$U = \left(\frac{\lambda_0 \alpha_0}{\beta_0}\right)^{1/2} (\mu - \mu_0) \sim t_{2\alpha_0}$$

It is then easy to show that (Theorem 8.6.3):

$$E(\mu) = \mu_0 \quad (\text{if } \alpha_0 > 1/2) \quad \text{and}$$

$$\mathrm{Var}(\mu) = \frac{\beta_0}{\lambda_0(\alpha_0 - 1)} \quad (\text{if } \alpha_0 > 1)$$

Marginal distribution of the mean

- Say we have done a Bayesian Analysis and end up with the Normal-Gamma posterior with parameters μ_1, λ_1, α_1 and β_1.
- We can then calculate the marginal posterior means of μ and τ

$$E(\mu|\mathbf{x}) = \mu_1 \quad (\text{if } \alpha_1 > 1/2) \quad \text{and} \quad E(\tau|\mathbf{x}) = \frac{\alpha_1}{\beta_1}$$

Can also obtain credible intervals for μ

- Find quantiles c_1 and c_2 such that $P(c_1 < U < c_2) = \gamma$
- Then

$$P\left(\mu_1 + c_1\left(\frac{\beta_1}{\lambda_1\alpha_1}\right)^{1/2} \leq \mu \leq \mu_1 + c_2\left(\frac{\beta_1}{\lambda_1\alpha_1}\right)^{1/2} \middle| \mathbf{x}\right) = \gamma$$

- E.g: for a 0.95% symmetric credible interval we set $c_1 = -T_{2\alpha_1}^{-1}(0.975)$ and $c_2 = T_{2\alpha_1}^{-1}(0.975)$

Example – Hotdogs (Exercise 8.5.7 in the book)

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

$$186, 181, 176, 149, 184, 190, 158, 139, 175, 148,$$
$$152, 111, 141, 153, 190, 157, 131, 149, 135, 132$$

Assume that these numbers are observed values from a random sample of twenty independent $N(\mu, \sigma^2)$ random variables, where μ and σ^2 are unknown.

$$\overline{x}_n = 156.85 \quad \text{and} \quad s_n^2 = \sum_{i=1}^{n}(x_i - \overline{x}_n)^2 = 9740.55$$

- Consider the Normal-Gamma prior for μ and τ with parameters $\mu_0 = 100$, $\lambda_0 = 3$, $\alpha_0 = 2$ and $\beta_0 = 2500$.
- Construct the apriori symmetric 95% credible interval for μ
- Find the posterior symmetric 95% credible interval for μ

Other priors

The usual improper prior for (μ, τ) is

$$p(\mu, \tau) = \frac{1}{\tau} \quad -\infty < \mu\infty, \ \tau > 0$$

- The posterior in this case is the Normal-Gamma distribution with parameters

$$\mu_1 = \overline{x}, \ \lambda_1 = n, \ \alpha_1 = (n-1)/2, \ \beta_1 = s_n^2/2$$

- Credible intervals turn out to be the same as confidence intervals (common for improper priors)

Other common options:

- μ and τ independent (a priori) with $\mu \sim N(\mu_0, \nu_0^2)$ and $\tau \sim \text{Gamma}(\alpha_0, \beta_0)$.
  - μ and τ will be dependent in the posterior.
- $\tau \sim \text{Gamma}(\alpha_0, \beta_0)$ but improper for μ: $p(\mu) = 1$