MVE051/MSG810 2017 Lecture 8

Petter Mostad

Chalmers

November 22, 2017

The Central Limit Theorem (CLT)

- Assume $X_1, \dots, X_n$ is a random sample from a distribution with expectation μ and variance σ^2. Then, when $n \to \infty$, we get that

$$\overline{X} \sim \text{Normal}(\mu, \sigma/\sqrt{n})$$

- For finite n the normal distribution can be used as an approximation. How large n needs to be for the approximation to be OK depends on what accuracy is needed, and on the properties of the distribution the X_i come from.
- Some distributions have no variance, then the CLT does not apply!

The normal distribution as an approximation

- It is the *mean value* $\overline{X}$ that becomes approximately normally distributed when the number of observations n increases. The sample itself does *not* become normally distributed just because n increases!
- However some random variables can be interpreted as a sum (or mean) of many (independent) random variables. Then, in some cases, they are well approximated by a normally distributed variable.
- Examples are:
  - The Binomial distribution with n large and p not too close to 0 or 1.
  - The Poisson distribution with a large intensity λ.
  - The Gamma distribution with a large α parameters.
  - The χ^2 distribution with many degrees of freedom.
- In such cases, the table for the Normal distribution can be used to compute approximate quantiles.

Confidence intervals

- A $100(1-\alpha)\%$ confidence interval for a parameter θ is a formula which from a random sample $X_1, \dots, X_n$ computes random variables L_1 and L_2 so that, for all θ,

$$\Pr\left[\theta \in [L_1, L_2]\right] = 1 - \alpha$$

- Correct interpretation of a confidence interval: If you generate N *new* random samples from the same distribution and from these compute N *new* confidence intervals according to the formula, then $100(1-\alpha)\%$ of these will contain θ as $N \to \infty$.
- WRONG interpreation of a confidence interval: Given a particular sample $x_1, \dots, x_n$, you know with 95% probaility that θ is in the confidence interval.
- However, if one interprets θ as a random variable and make certain assumptions, the second interpretation can become correct. These assumptions are often reasonable. This underlies the popularity of the confidence interval in applications.

Example 1: Confidence interval for μ in a Normal(μ, σ) distribution

- If $X_1, \dots, X_n \sim$ Normal(μ, σ) is a random sample, if $z_{\alpha/2}$ is so that $[-z_{\alpha/2}, z_{\alpha/2}]$ contains $100(1-\alpha)\%$ of the probability in a standard normal distribution, and if we define

$$\begin{aligned} L_1 &= \overline{X} - z_{\alpha/2}\sigma/\sqrt{n} \\ L_2 &= \overline{X} + z_{\alpha/2}\sigma/\sqrt{n} \end{aligned}$$

  then $[L_1, L_2]$ is a $100(1-\alpha)\%$ confidence interval for μ.
- Note how we find $z_{\alpha/2}$ in the table for the standard normal distribution. Traditionally we use $\alpha = 0.05$, giving $z_{0.05/2} = 1.96$.
- Note: The formulas for L_1 and L_2 contain σ, so this interval can only be used if σ^2, the variance of the distribution, is known. (It is not enough to compute the sample variance of the data).
- The proof is based on using that $\overline{X} \sim$ Normal$(\mu, \sigma/\sqrt{n})$.

The distribution for the variance estimator $S^2 = \hat{\sigma}^2$

- The confidence interval for μ was constructed based on knowing the distribution for the estimator $\overline{X}$ for μ. In the same way we base a confidence interval for σ^2 on the estimator $\hat{\sigma}^2$.
- $X_1, \dots, X_n \sim \text{Normal}(\mu, \sigma)$ and we define the estimator

  $$S^2 = \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

  the the distribution of this estimator satisfies

  $$(n-1)S^2/\sigma^2 \sim \chi^2(n-1),$$

  i.e., $(n-1)S^2/\sigma^2$ has a χ^2 distribution with $n-1$ degrees of freedom.
- A proof can be constructed by (see Milton appendix C)
  - first showing that S^2 och $\overline{X}$ är independent random variables (e.g., use moment generating functions).
  - then using this to compute the moment-generating function for $(n-1)S^2/\sigma^2$ and showing that it corresponds to that of the $\chi^2(n-1)$ distribution.

Example 2: Confidence interval for σ^2

- If $X_1, \dots, X_n \sim \text{Normal}(\mu, \sigma)$ is a random sample, if $\chi^2_{n-1,\alpha/2}$ and $\chi^2_{n-1,1-\alpha/2}$ are so that $[\chi^2_{n-1,\alpha/2}, \chi^2_{n-1,1-\alpha/2}]$ contains $100(1-\alpha)\%$ of the probability in a $\chi^2(n-1)$ distribution, and if we define

$$\begin{aligned} L_1 &= (n-1)S^2/\chi^2_{n-1,1-\alpha/2} \\ L_2 &= (n-1)S^2/\chi^2_{n-1,\alpha/2} \end{aligned}$$

  then $[L_1, L_2]$ is a $100(1-\alpha)\%$ confidence interval for σ^2.
- Note how we find $\chi^2_{n-1,1-\alpha/2}$ and $\chi^2_{n-1,\alpha/2}$ in the table for the $\chi^2(k)$ distribution.
- Note: The formulas for L_1 och L_2 do not contain μ so this interval can be used even when μ is unknown.
- The proof is based on using that $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.

The (Student) t distribution

- The random variable X has a (Student) t distribution with γ degrees of freedom, we write $X \sim T(\gamma)$, if the density is

$$f(x) = \frac{\Gamma(\gamma+1)/2}{\Gamma(\gamma/2)\sqrt{\pi\gamma}}\left(1+\frac{x^2}{\gamma}\right)^{-(\gamma+1)/2}$$

- When $\gamma \to \infty$ the t distribution will approach a standard normal distribution. When γ is smaller, the density is more pointy at the center, and has "heavier tails", than the standard normal.
- We have $\mathbb{E}[X] = 0$ (if $\gamma \leq 1$ the expectation does not exist) and $\mathrm{Var}[X] = \gamma/(\gamma-2)$ (if $\gamma \leq 2$ the variance does not exist).
- An important property: If $Z \sim \mathrm{Normal}(0,1)$ and $X \sim \chi^2(\gamma)$ are independent, then

$$\frac{Z}{\sqrt{X/\gamma}} \sim T(\gamma).$$

- Tables for the Student t distribution for various γ values are available in Milton.

The distribution for $(\overline{X} - \mu)/(S/\sqrt{n})$

- Our earlier confidence interval for μ depended on σ. We now construct a confidence interval for μ that depends on S^2 instead. We do this by studying a function of $\overline{X}$ och S^2.
- If $X_1, \ldots, X_n \sim \text{Normal}(\mu, \sigma)$ is a random sample then

$$(\overline{X} - \mu)/(S/\sqrt{n}) \sim T(n-1)$$

  In other words, the statistic has a t distribution with $n-1$ degrees of freedom.
- A proof can be based on
  - $\overline{X} \sim \text{Normal}(\mu, \sigma/\sqrt{n})$.
  - $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
  - $\overline{X}$ and S^2 are independent.
  - The property of the t distribution mentioned on the previous overhead.

Example 3: A confidence interval for μ not depending on σ

- If $X_1, \dots, X_n \sim \text{Normal}(\mu, \sigma)$ is a random sample, if $t_{\alpha/2}$ is so that $[-t_{\alpha/2}, t_{\alpha/2}]$ contains $100(1-\alpha)\%$ of the probability in a t distribution with $n-1$ degrees of freedom, and if we define

$$\begin{aligned} L_1 &= \overline{X} - t_{\alpha/2} S/\sqrt{n} \\ L_2 &= \overline{X} + t_{\alpha/2} S/\sqrt{n} \end{aligned}$$

  then $[L_1, L_2]$ is a $100(1-\alpha)\%$ confidence interval for μ.
- Note how we find $t_{\alpha/2}$ in the table for the t distribution.
- The formulas for L_1 and L_2 do not contain σ, so this interval *can* be used if the distribution variance σ^2 is unknown.
- The proof is based on using that $(\overline{X} - \mu)/(S/\sqrt{n}) \sim \text{T}(n-1)$.

Example 4: Approximate confidence interval for μ based on CLT and known σ

- If $X_1, \ldots, X_n$ is a random sample from a distribution with expectation μ and variance σ^2, if $z_{\alpha/2}$ is so that $[-z_{\alpha/2}, z_{\alpha/2}]$ contains $100(1-\alpha)\%$ of the probability in a standard normal distribution, and if we define

$$\begin{aligned} L_1 &= \overline{X} - z_{\alpha/2}\sigma/\sqrt{n} \\ L_2 &= \overline{X} + z_{\alpha/2}\sigma/\sqrt{n} \end{aligned}$$

  then $[L_1, L_2]$ is an *approximate* $100(1-\alpha)\%$ confidence interval for μ *if n is large*.
- Note: To use this, the variance σ^2 must exist.
- The proof uses that, for large n we have, approximately, $\overline{X} \sim \text{Normal}(\mu, \sigma/\sqrt{n})$.

Example 5: Approximate confidence interval for μ based on CLT

- If $X_1, \dots, X_n$ is a random sample from a distribution with expectation μ and variance σ^2, if $z_{\alpha/2}$ is so that $[-z_{\alpha/2}, z_{\alpha/2}]$ contains $100(1-\alpha)\%$ of the probability in a standard normal distribution, and if we define

$$\begin{aligned} L_1 &= \overline{X} - z_{\alpha/2} S/\sqrt{n} \\ L_2 &= \overline{X} + z_{\alpha/2} S/\sqrt{n} \end{aligned}$$

  where S is the sample standard deviation, then $[L_1, L_2]$ is an *approximate* $100(1-\alpha)\%$ confidence interval for μ *if n is large*.
- Motivation: For large n we have, approximately, $\overline{X} \sim \text{Normal}(\mu, \sigma/\sqrt{n})$ and also $S \approx \sigma$.
- One may use $t_{\alpha/2}$ instead of $z_{\alpha/2}$, but the difference is small as n is large.
- Note: For this to hold, the variance σ^2 must exist. One can always compute S from a sample, that S exists does not imply that σ exists!