Complications in Using MLE

- Maximum occurs at a discontinuous point.

**Example:** Suppose $y_1, \ldots, y_n \sim \text{Unif}(0, \theta)$. Find the MLE of $\theta$.

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} 1_{\{0 \leq y_i \leq \theta\}}$$

$$= \frac{1}{\theta^{n}} 1_{\{\min(y_i) \geq 0\}} 1_{\{\max(y_i) \leq \theta\}}$$

$$= \begin{cases} 
0 & \theta < \max(y_i) \\
\frac{1}{\theta^{n}} & \theta \geq \max(y_i) 
\end{cases}$$

So $\hat{\theta} = \arg\max_{\theta} L(\theta) = \max(y_i)$.

$\hat{\theta} = \max y_i$ is biased. Why?

- Close-form solution does not exist

**Example:** Suppose $y_1, \ldots, y_n \sim \text{Gamma}(\alpha, \beta)$. Find the MLE.
Confidence Interval

- **Aim:** How to use the sample to calculate two numbers that define an interval that will enclose the unknown parameter with certain probability (confidence).

- The resulting random interval is called a confidence interval.

- The probability that the interval contains the unknown parameter is called its confidence coefficient.

\[
P(\text{LCL} \leq \theta \leq \text{UCL}) = 1 - \alpha
\]

$\alpha$ is usually small, for example, 5% or 2.5%.

**LCL**$(y_1, \ldots, y_n)$: lower confidence limit

**UCL**$(y_1, \ldots, y_n)$: upper confidence limit
Case 1: Normal with Known Variance

Suppose $\mu \sim N(\mu, \sigma^2)$ with $\sigma^2$ known.

(1) Define (pivotal statistic)

$$z = \frac{\hat{\mu} - \mu}{\sigma} \sim N(0, 1).$$

(2) Locate values $z_{\alpha/2}$ and $-z_{\alpha/2}$ that place a probability of $\alpha/2$ in each tail of $N(0, 1)$. For example, $z_{0.025} = 1.96$.

$$1 - \alpha = P(-z_{\alpha/2} \leq z \leq z_{\alpha/2})$$

$$= P(-z_{\alpha/2} \leq \frac{\hat{\mu} - \mu}{\sigma} \leq z_{\alpha/2})$$

$$= P(-z_{\alpha/2}\sigma \leq \hat{\mu} - \mu \leq z_{\alpha/2}\sigma)$$

$$= P(\hat{\mu} - z_{\alpha/2}\sigma \leq \mu \leq \hat{\mu} + z_{\alpha/2}\sigma)$$

**Theorem 8.2** Let $\mu \sim N(\mu, \sigma^2)$. Then a $(1 - \alpha)100\%$ confidence interval for $\mu$ is

$$\hat{\mu} - z_{\alpha/2}\sigma \text{ to } \hat{\mu} + z_{\alpha/2}\sigma$$
Example: ASTM Standard E23 defines standard test methods for notched bar impact testing of metallic materials. The Charpy V-notch (CVN) technique measures impact energy and is often used to determine whether or not a material experiences a ductile-to-brittle transition with decreasing temperature. Ten measurements of impact energy ($J$) on specimens of A238 steel cut at $60^0$C are as follows: 64.1, 64.7, 64.5, 64.6, 64.5, 64.3, 64.6, 64.8, 64.2 and 64.3. Assume that impact energy is normally distributed with $\sigma = 1J$.

Find a 95% CI for $\mu$, the mean impact energy.

$$\bar{y} - z_{\alpha/2}\sigma_{\bar{y}} \leq \mu \leq \bar{y} + z_{\alpha/2}\sigma_{\bar{y}}$$

$$\bar{y} - 1.96\frac{1}{\sqrt{10}} \leq \mu \leq \bar{y} + 1.96\frac{1}{\sqrt{10}}$$

$$64.46 - 1.96\frac{1}{\sqrt{10}} \leq \mu \leq 64.46 + 1.96\frac{1}{\sqrt{10}}$$

$$63.84 \leq \mu \leq 65.08$$

How many specimens must be tested to ensure that the 95% CI of $\mu$ has a length of at most $1.0J$?

$$n = \left[ \frac{(1.96)1}{0.5} \right]^2 = 15.37.$$
**Interpreting a CI**

- Can we conclude: The true mean $\mu$ is within the interval (63.84, 65.08) with probability 0.95? – **NO**

- The statement $63.84 \leq \mu \leq 65.08$ is either correct (true with probability 1) or incorrect (false with probability 1).

- Remember that a CI is a **random interval** and the correct interpretation of a $100(1 - \alpha)\%$ CI should depend on the relative frequency view of probability.

- We conclude:

  If we were to repeatedly collect a sample of size $n$ and construct a 95% CI for each sample, then we expect 95% of the intervals to enclose the true parameter $\mu$. 
m=20; n=10;
y = normrnd(0, 0.5, m,n);
y_mean = mean(y,2);
e = norminv(0.95)*0.5/sqrt(n)*ones(m,1)
errorbar(1:m, y_mean, e);
h = line([1, m], [0, 0]);
set(h, 'color', [1 0 0]);

LCL = y_mean - e;
UCL = y_mean + e;
sum(LCL > 0) + sum(UCL < 0)

ans =

2
• If $\bar{y}$ is the sample mean of a random sample of size $n$ from $N(\mu, \sigma^2)$, the $(1 - \alpha)100\%$ CI is

$$\bar{y} - z_{\alpha/2} \sigma / \sqrt{n} \leq \mu \leq \bar{y} + z_{\alpha/2} \sigma / \sqrt{n}$$

• The length of the CI is equal to $2z_{\alpha/2} \sigma / \sqrt{n}$.

• What’s the relationship between the length of a CI and
  
  • the confidence coefficient $(1 - \alpha)100\%$?
  • the sample size $n$?

  **Q:** How many sample size we should choose in order to have the length of CI less than $l_0$.

  $$2z_{\alpha/2} \sigma / \sqrt{n} \leq l_0$$

  $$n \geq \left( \frac{z_{\alpha/2} \sigma}{l_0/2} \right)^2$$
Sampling Dist Related to Normal

A random sample \( y_1, y_2, \ldots, y_n \) is drawn from \( N(\mu, \sigma^2) \).

\[
\text{sample mean } \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \\
\text{sample var } s^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n - 1}
\]

- Recall that \( \chi^2(\nu) = z_1^2 + z_2^2 + \cdots + z_{\nu}^2 \), where each \( z_i \sim N(0, 1) \).

- \( (n - 1)s^2/\sigma^2 \sim \chi^2(n - 1) \)

- Let \( z \) be a standard normal and \( \chi^2 \) be a chi-square with \( \nu \) degrees of freedom. If \( z \) and \( \chi^2 \) are independent, then

\[
t = \frac{z}{\sqrt{\chi^2/\nu}}
\]

has a **Student’s t distribution** with \( \nu \) degree of freedom.
Case 2: Normal with Unknown Variance

(Example 8.6) Let $\bar{y}$ and $s^2$ be the sample mean and variance based on a random sample of $n$ normal$(\mu, \sigma^2)$ observations

Start with the pivotal statistic

$$t = \frac{\bar{y} - \mu}{s/\sqrt{n}} = \frac{\bar{y} - \mu}{\sigma/\sqrt{n}} \sqrt{\frac{(n-1)s^2}{\sigma^2}} / (n - 1)$$

$\sim$ Student’s t distribution

$$1 - \alpha = P(-t_{\alpha/2,n-1} \leq t \leq t_{\alpha/2,n-1})$$

$$= P(-t_{\alpha/2,n-1} \leq \frac{\bar{y} - \mu}{s/\sqrt{n}} \leq t_{\alpha/2,n-1})$$

$$= P(\bar{y} - t_{\alpha/2,n-1} \left( \frac{s}{\sqrt{n}} \right) \leq \mu \leq \bar{y} + t_{\alpha/2,n-1} \left( \frac{s}{\sqrt{n}} \right))$$

Take a look of Example 8.7