Econ 325
Notes on Point Estimator and Confidence Interval[1]
By Hiro Kasahara

Point Estimator

Parameter, Estimator, and Estimate

The normal probability density function is fully characterized by two constants: population mean μ and population variance σ^2. The probability mass function of Bernoulli random variable is fully defined by the population fraction of "success", p. These constants are called *parameters* and we generally use the Greek letter θ to denote them.

We are often interested in knowing the population parameter such as population mean and population variance. To guess the population value of mean and variance, we use their sample analogues, i.e., sample mean and sample variance.

A point *estimator* of θ is a function of the random sample, denoted by $\hat{\theta}$:

$$\hat{\theta} = \hat{\theta}(X_1, X_2, ..., X_n).$$

Here, the right hand side of the equation provides a mapping from the sample $\{X_1, X_2, ..., X_n\}$ to real value. Namely, $\hat{\theta}(X_1, X_2, ..., X_n)$ a "formula" to compute the sample analog of corresponding population parameter (e.g., for sample mean $\bar{X}$, we have $\hat{\theta}(X_1, X_2, ..., X_n) = \frac{1}{n}\sum_{i=1}^n X_i$). The estimator $\hat{\theta}$ is a random variable because the sample $\{X_1, X_2, ..., X_n\}$ is randomly drawn.

When we evaluate $\hat{\theta}(X_1, X_2, ..., X_n)$ at the realized sample, then $\hat{\theta}$ is called an *estimate*. The evaluated value of the function $\hat{\theta}(X_1, X_2, ..., X_n)$ at the realized sample is not a random variable any more—rather, it is constant.

Unbiasedness

An estimator $\hat{\theta}$ is said to be an *unbiased estimator of the parameter* θ if

$$E[\hat{\theta}] = \theta.$$

The *bias* of an estimator $\hat{\theta}$ is defined as

$$\text{Bias} = E[\hat{\theta}] - \theta.$$

The bias of an unbiased estimator is zero by definition.

Example 1. *The sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is an unbiased estimator of the population mean μ because $E[\bar{X}] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}\sum_{i=1}^n \mu = \mu$. The sample variance $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ is an unbiased estimator of the population variance σ^2. However, the estimator $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ is not an unbiased estimator of σ^2 because $E[\hat{\sigma}^2] = E[\frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2] = \frac{n-1}{n}E[\frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2] = \frac{n-1}{n}\sigma^2 < \sigma^2$.*

Example 2. *Given a random sample of size n $\{X_1, X_2, ..., X_n\}$, consider an estimator $\hat{X} = X_1$, which only uses the first observation while ignores all other $n-1$ observations. This estimator is an unbiased estimator of μ because $E[X_1] = \mu$. We can also consider an estimator defined by the weighted average of X_i's as $\hat{X} = \sum_{i=1}^n w_i X_i$, where $\{w_i\}_{i=1}^n$ is a sequence of n numbers such that $\sum_{i=1}^n w_i = 1$. Then, this estimator is an unbiased estimator of μ because $E[\sum_{i=1}^n w_i X_i] = \sum_{i=1}^n w_i \mu = \mu$.*

Example 3. *The sample fraction $\hat{p}$ is an unbiased estimator fo the population fraction p. This is because the sample fraction is viewed as the sample average of n independent Bernoulli random variables, i.e., $\hat{p} = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, where $X_i = 0$ with probability $1-p$ and $X_i = 1$ with probability p. Taking the expectation, $E[\hat{p}] = \frac{1}{n}\sum_{i=1}^n E[X_i] = \frac{1}{n}\sum_{i=1}^n p = p$.*

Efficiency

Consider the case for $n = 2$ and let X_1 and X_2 are randomly sampled from the population distribution with mean μ and variance σ^2. Consider the following two estimators for μ:

$$\bar{X} = \frac{1}{2}(X_1 + X_2) \quad \text{and} \quad \tilde{X} = \frac{1}{3}X_1 + \frac{2}{3}X_2.$$

Both $\bar{X}$ and $\tilde{X}$ are unbiased estimators because $E[\bar{X}] = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$ and $E[\tilde{X}] = \frac{1}{3}\mu + \frac{2}{3}\mu = \mu$. In fact, we may consider an estimator of the form given by $aX_1 + (1-a)X_2$ for any fixed value of a and we can verify that $aX_1 + (1-a)X_2$ is an unbiased estimator because $E[aX_1 + (1-a)X_2] = \mu$.

While unbiasedness is a desirable property of estimators, we have multiple unbiased estimators. Which estimators do we want to choose among all unbiased estimators? The answer is: the estimator that has the smallest variance. Intuitively, the smaller the variance is, the closer the realized value of the estimator is to the population mean on average. In fact, if the variance of the unbiased estimator is zero, we have population mean for every realized value of the estimator.

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two unbiased estimators. Then $\hat{\theta}_1$ is said to be *more efficient* than $\hat{\theta}_2$ if

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2).$$

If $\hat{\theta}_1$ is an unbiased estimator that has the smallest variance among all unbiased estimators, then $\hat{\theta}_1$ is said to be *the most efficient, or the minimum variance unbiased estimator.*

Example 4. *Consider the case for $n = 2$ and X_1 and X_2 are randomly sampled from the population distribution with mean μ and variance σ^2. What is the most efficient unbiased estimator? To answer this, we consider the class of unbiased estimator of the form $aX_1 + (1-a)X_2$ for any fixed value a. The variance of $aX_1 + (1-a)X_2$ is given by*

$$Var(aX_1 + (1-a)X_2) = \{a^2 + (1-a)^2\}\sigma^2,$$

where $Cov(X_1, X_2) = 0$ by random sampling. Therefore, we may find the most efficient unbiased estimator by minimizing $g(a) = a^2 + (1-a)^2 = 2a^2 - 2a + 1$ with respect to a. The first order condition is given by $g'(a) = 4a - 2 = 0$ so that $g(a)$ is minimized at $a = 1/2$. Therefore, $\frac{1}{2}X_1 + \frac{1}{2}X_2 = \bar{X}$ is the most efficient unbiased estimator for μ.

Consistency

A point estimator $\hat{\theta}$ is said to be *consistent* if $\hat{\theta}$ converges in probability to θ, i.e., for every $\epsilon > 0$, $\lim_{n\to\infty} P(|\hat{\theta} - \theta| < \epsilon) = 1$ (see Law of Large Number).

Example 5. *Suppose that X_1, X_2, ..., X_n are randomly sampled from a population with mean μ and variance σ^2. Is $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ an consistent estimator of μ? How about $\frac{1}{n-1}\sum_{i=1}^n X_i$ and $\frac{1}{n-1}\sum_{i=1}^{n-1} X_i$? Are these two estimators consistent?*

The sample variance $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ is a consistent estimator of σ^2. Is the estimator $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ an consistent estimator of σ^2?

Example 6. *Suppose that X_1, X_2, ..., X_n are n independent Bernoulli random variables, i.e., $\hat{p} = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, where $X_i = 0$ with probability $1-p$ and $X_i = 1$ with probability p. The parameter p is the population fraction of individuals with $X_i = 1$. The sample fraction is defined as $\hat{p} = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. By the Law of Large Numbers, the sample fraction $\hat{p}$ is a consistent estimator of the population fraction p.*

The variance of $\hat{p}$ is given by $Var(\hat{p}) = Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \left(\frac{1}{n}\right)^2 Var(X_1 + X_2 + ... + X_n) = \frac{n\times Var(X_i)}{n^2} = \frac{Var(X_i)}{n} = \frac{p(1-p)}{n}$, where the last line follows from $Var(X_i) = E[(X_i - p)^2] = (0-p)^2 \times (1-p) + (1-p)^2 \times p = p(1-p)$. Because $Var(\hat{p}) = \frac{p(1-p)}{n}$ involves the unknown population parameter p, we do not know the value of $Var(\hat{p}) = \frac{p(1-p)}{n}$. We can construct an estimator for $Var(\hat{p}) = \frac{p(1-p)}{n}$ by replacing p with $\hat{p}$, where the latter can be computed from the sample, as $\widehat{Var(\hat{p})} = \frac{\hat{p}(1-\hat{p})}{n}$. $\widehat{Var(\hat{p})} = \frac{\hat{p}(1-\hat{p})}{n}$ is a consistent estimator of $Var(\hat{p})$.

Confidence Interval

We may estimate interval rather than a point. The idea of interval estimation is to construct a *random* interval such that the constructed interval contains the true parameter θ with a pre-specified probability, $1-\alpha$. Such an interval is called $(1-\alpha)$ *percent confidence interval*, where $1-\alpha$ is called the *confidence level*. The confidence interval is characterized by the lower limit (L) and the upper limit (U), both of which is a function of the random sample X_1, ..., X_n so that

$$P\left(L(X_1, ..., X_n) \le \theta \le U(X_1, ..., X_n)\right) = 1 - \alpha.$$

Note that both $L(X_1, ..., X_n)$ and $U(X_1, ..., X_n)$ are random variables.

The case that n is large or $\hat{\theta} \sim N(\theta, Var(\hat{\theta}))$ with $Var(\hat{\theta})$ known.

Suppose that a point estimator $\hat{\theta}$ is approximately normally distributed with mean θ and variance $Var(\hat{\theta})$, i.e.,

$$\hat{\theta} \sim N\left(\theta, Var(\hat{\theta})\right).$$

Two representative cases are:

1. X_1, X_2, ..., X_n are randomly sampled from some distribution that is different from normal distribution but the sample size n is large. In this case, $\hat{\theta}$ is defined as the

average of random variable, i.e., $\hat{\theta} = \bar{X} = \frac{1}{n}X_i$ so that we may apply the Central Limit Theorem to have $\hat{\theta} \sim N(E[X_i], Var(X_i)/n)$; for example, the sample mean $\bar{X} \sim N(\mu, \sigma^2/n)$. When n is large, we may essentially treat $Var(X_i)$ as if it is known and given by the sample variance.

2. X_1, X_2, ..., X_n are randomly sampled from normal distribution $N(\mu, \sigma^2)$ with known variance σ^2. In this case, the sample average is normally distributed with mean μ and variance σ^2/n.

In these cases, we may construct the 95 percent confidence interval with

$$[L, U] = \left[\hat{\theta} - 1.96\sqrt{Var(\hat{\theta})},\ \hat{\theta} + 1.96\sqrt{Var(\hat{\theta})}\right],$$

so that

$$P\left(\hat{\theta} - 1.96\sqrt{Var(\hat{\theta})} \leq \theta \leq \hat{\theta} + 1.96\sqrt{Var(\hat{\theta})}\right) = 0.95.$$

In general, the confidence interval with confidence level $(1-\alpha)$ is constructed as

$$P\left(\hat{\theta} - z_{\alpha/2}\sqrt{Var(\hat{\theta})} \leq \theta \leq \hat{\theta} + z_{\alpha/2}\sqrt{Var(\hat{\theta})}\right) = 1 - \alpha, \tag{1}$$

where $z_{\alpha/2}$ is determined such that $P(Z \geq z_{\alpha/2}) = \alpha/2$ when $Z \sim N(0,1)$. Here, $z_{\alpha/2}\sqrt{Var(\hat{\theta})}$ is called as the *margin of error*.

We may confirm (1) by reformulating the inequality on the left hand side of (1) in terms of a standardized random variable $\frac{\hat{\theta}-\theta}{\sqrt{Var(\hat{\theta})}}$ as follows.

$$\begin{aligned}
&P\left(\hat{\theta} - z_{\alpha/2}\sqrt{Var(\hat{\theta})} \leq \theta \leq \hat{\theta} + z_{\alpha/2}\sqrt{Var(\hat{\theta})}\right) \\
=&P\left(\left\{\hat{\theta} - \theta \leq z_{\alpha/2}\sqrt{Var(\hat{\theta})}\right\} \text{ and } \left\{z_{\alpha/2}\sqrt{Var(\hat{\theta})} \leq \hat{\theta} - \theta\right\}\right) \\
=&P\left(\left\{\frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} \leq z_{\alpha/2}\right\} \text{ and } \left\{-z_{\alpha/2} \leq \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}}\right\}\right) \\
=&P\left(z_{\alpha/2} \leq \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} \leq z_{\alpha/2}\right) \quad \text{where } Z = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}} \sim N(0,1) \qquad (2) \\
=&P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha.
\end{aligned}$$

Example 7 (Confidence Interval for population mean μ)**.** *Given the random sample X_1, ..., X_n drawn from $N(\mu, \sigma^2)$ and σ^2 is known, the sample average $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ is an estimator of μ with $\sqrt{Var(\bar{X})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$. Therefore, 95 percent confidence interval for μ is given by*

$$[L, U] = \left[\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right].$$

If we would like to construct 90 percent confidence interval with $\alpha = 0.1$, $z_{0.05} = 1.645$ *(i.e.,* $P(Z \geq 1.645) = 0.05$ *for* $Z \sim N(0,1)$*). Therefore, 95 percent confidence interval for* μ *is given by*

$$[L, U] = \left[\bar{X} - 1.645 \frac{\sigma}{\sqrt{n}}, \bar{X} + 1.645 \frac{\sigma}{\sqrt{n}} \right].$$

Example 8 (Survey on the U.S. presidential election in Florida)**.** *The survey was conducted between Oct. 20 and 24, 2016, in Florida after the third and final presidential debate. The survey result shows that, among 1166 likely registered voters who support either Clinton or Trump, there are 602 Clinton voters and 564 Trump voters. What is the 95 precent confidence interval for the population fraction of Clinton voters?*

Let p *be the population fraction of Clinton voters. Each voter's preference is a Bernoulli random variable* X_i *with* $P(X_i = 1) = p$ *and* $P(X_i = 0) = 1 - p$*, where* $X_i = 1$ *means Clinton voter while* $X_i = 0$ *means Trump voter. The sample average is given by* $\hat{p} = 0.516$*. The standard deviation of* $\hat{p}$ *is given by* $\sqrt{p(1-p)/n}$*, which can be estimated as* $\sqrt{\hat{p}(1-\hat{p})/n} = \sqrt{0.516(1-0.516)/1166} = 0.01463$*. The margin of error is, therefore,* $1.96 \times 0.01463 = 0.0287$*. Then, we may construct the 95 percent confidence interval with*

$$[L, U] = \left[\hat{p} - 1.96\sqrt{\hat{p}(1-\hat{p})/n},\ \hat{p} + 1.96\sqrt{\hat{p}(1-\hat{p})/n} \right] = [0.488, 0.545].$$

Therefore, the population fraction of Clinton voters is between 0.488 *and* 0.545 *with probability 95 percent. Therefore, in this Florida's poll, Clinton's lead is "within the margin of error".*

Example 9 (Survey on the U.S. presidential election in North Carolina)**.** *The survey was conducted between November 3 and 6, 2016, in North Carolina. The survey result shows that, among 791 likely registered voters who support either Clinton or Trump, there are 400 Clinton voters and 391 Trump voters. What is the 95 precent confidence interval for the population fraction of Clinton voters?*

The knowledge of $Var(\hat{\theta})$ is required for constructing confidence interval as shown in (1). Typically, $Var(\hat{\theta})$ depends on population parameter that is unknown (e.g., $Var(\bar{X}) = \sigma^2/n$ and $Var(\hat{p}) = p(1-p)/n$) but we can estimate $Var(\hat{\theta})$. The estimator of $Var(\bar{X})$ and $Var(\hat{p})$ are given by

$$\hat{Var}(\bar{X}) = s^2/n \quad \text{and} \quad \hat{Var}(\hat{p}) = \hat{p}(1-\hat{p})/n.$$

When $Var(\hat{\theta})$ is not known, we replace $Var(\hat{\theta})$ with its estimator $\hat{Var}(\hat{\theta})$ in constructing confidence interval. In the above example of the U.S. presidential election (Example 8), this is what we did: we replaced $Var(\hat{p}) = \frac{p(1-p)}{n}$ with its estimator $\frac{\hat{p}(1-\hat{p})}{n}$. This is fine as long as the sample size n is large because the estimator of $Var(\hat{p})$ converges in probability to $Var(\hat{p})$ and we may essentially treat $Var(\hat{p})$ as known in constructing the confidence interval. When n is small, however, this is not the case anymore. The randomness of the estimator of $Var(\hat{p})$ does not go away when n is small and the constructed confidence interval in (1) by replacing $Var(\hat{\theta})$ with its estimator does not contain θ with probability $(1-\alpha)$ anymore.

The case that n is small

When n is small, it is generally difficult to construct confidence interval for two reasons. First, we may not use the Central Limit Theorem to claim that $\hat{\theta}$ is normally distributed. Second, replacing $Var(\hat{\theta})$ with its estimator $\hat{Var}(\hat{\theta})$ introduces an additional source of randomness.

In both cases, the standardized random variable using the estimator of $Var(\hat{\theta})$

$$\frac{\hat{\theta} - \theta}{\sqrt{\hat{Var}(\hat{\theta})}}$$

is not a standard normal random variable, and therefore the confidence interval (1) is not valid any more because (1) is constructed under the assumption that $\frac{\hat{\theta}-\theta}{\sqrt{Var(\hat{\theta})}} \sim N(0,1)$ (see (2)).

While it is generally difficult to construct confidence interval, there is one exceptional case where we may construct confidence interval using Student's t-distribution.

Suppose that we have a random sample X_1, X_2, ..., X_n from $N(\mu, \sigma^2)$. In this case, we have

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim \text{Student's t distribution with } (n-1) \text{ degrees of freedom.}$$

Therefore, the confidence interval for μ with confidence level $(1-\alpha)$ is constructed as

$$P\left(\bar{X} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right) = 1 - \alpha, \quad (3)$$

where $t_{n-1,\alpha/2}$ is determined such that $P(T \geq t_{n-1,\alpha/2}) = \alpha/2$ when $T \sim$ Student's t distribution with $(n-1)$ degrees of freedom. We may confirm (3) by reformulating the inequality on the left hand side of (3) in terms of a standardized random variable $\frac{\bar{X}-\mu}{s/\sqrt{n}}$ as follows.

$$\begin{aligned}
&P\left(\bar{X} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right) \\
=&P\left(\left\{\frac{\bar{X} - \mu}{s/\sqrt{n}} \leq t_{n-1,\alpha/2}\right\} \text{ and } \left\{-t_{n-1,\alpha/2} \leq \frac{\bar{X} - \mu}{s/\sqrt{n}}\right\}\right) \\
=&P\left(-t_{n-1,\alpha/2} \leq \frac{\bar{X} - \mu}{s/\sqrt{n}} \leq t_{n-1,\alpha/2}\right) \\
=&P(-t_{n-1,\alpha/2} \leq T \leq t_{n-1,\alpha/2}) = 1 - \alpha,
\end{aligned} \quad (4)$$

where $T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim$ Student's t distribution with $(n-1)$ degrees of freedom in (4).

A few comments. First, we need the assumption that X_1, X_2, ..., X_n are drawn from the normal distribution. If X_i is a Bernoulli random variable, then $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim$ does not follow t-distribution. Second, as $n \to \infty$, $s^2 \to_p \sigma^2$, so that the Student's t distribution converges to the standard normal distribution as $n \to \infty$. In fact, at $n = 31$, the critical value for 95 percent confidence interval using t-distribution is given by 2.042 which is close to 1.96.

Confidence interval for sample variance

Suppose that $\{X_1, X_2, ..., X_n\}$ is a random sample from a normal distribution with $E[X_i] = \mu$ and $\text{Var}[X_i] = \sigma^2$. Then, the random variable $\frac{(n-1)s_n^2}{\sigma^2}$ has a distribution known as the chi-square distribution with $n-1$ degree of freedom which we denote by χ^2_{n-1}, i.e.,

$$\frac{(n-1)s_n^2}{\sigma^2} = \chi^2_{n-1}. \quad (5)$$

Let $\chi^2_{n-1,\alpha/2}$ and $\chi^2_{n-1,1-\alpha/2}$ be the value such that $P(\chi^2_{n-1} > \chi^2_{n-1,\alpha/2}) = \alpha/2$ and $P(\chi^2_{n-1} > \chi^2_{n-1,1-\alpha/2}) = 1 - \alpha/2$ so that

$$P\left(\chi^2_{n-1,1-\alpha/2} < \frac{(n-1)s_n^2}{\sigma^2} < \chi^2_{n-1,\alpha/2}\right) = 1 - \alpha.$$

Then, we may construct the confidence interval for σ^2 from the sample variance $s_n = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ as follows.

$$\begin{aligned} 1 - \alpha &= P\left(\chi^2_{n-1,1-\alpha/2} < \frac{(n-1)s_n^2}{\sigma^2} < \chi^2_{n-1,\alpha/2}\right) \\ &= P\left(\frac{1}{\chi^2_{n-1,\alpha/2}} < \frac{\sigma^2}{(n-1)s_n^2} < \frac{1}{\chi^2_{n-1,1-\alpha/2}}\right) \\ &= P\left(\frac{(n-1)s_n^2}{\chi^2_{n-1,\alpha/2}} < \sigma^2 < \frac{(n-1)s_n^2}{\chi^2_{n-1,1-\alpha/2}}\right). \end{aligned}$$

Therefore,

$$P\left(L < \sigma^2 < U\right) 1 = 1 - \alpha$$

with

$$L = \frac{(n-1)s_n^2}{\chi^2_{n-1,\alpha/2}} \quad \text{and} \quad U = \frac{(n-1)s_n^2}{\chi^2_{n-1,1-\alpha/2}}.$$

Example 10 (Confidence Interval and Hypothesis Testing for Sample Variance)**.** *Suppose that you are a plant manager for producing electrical devices operated by a thermostatic control. According to the engineering specifications, the standard deviation of the temperature at which these controls actually operate should not exceed 2.0 degrees Fahrenheit. As a plant manager, you would like to know how large the (population) standard deviation* σ *is. We assume that the temperature is normally distributed. Suppose that you randomly sampled 25 of these controls, and the sample variance of operating temperatures was* $s_n^2 = 2.36$ *degrees Fahrenheit. (i) Compute the 95 percent confidence interval for the population standard deviation* σ*. (ii) Test the null hypothesis* $H_0 : \sigma = 2$ *against the alternative hypothesis* $H_1 : \sigma > 2$ *at the significance level* $\alpha = 0.05$*.*

The distribution of $\frac{(n-1)s_n^2}{\sigma^2}$ *is given by chi-square distribution with* $(n-1)$ *degrees of freedom. Let* χ^2_{n-1} *be a random variable distributed by chi-square distribution with* $(n-1)$ *degrees of freedom and let* $\chi^2_{n-1,\alpha}$ *be the value such that* $\Pr(\chi^2_{n-1} < \chi_{n-1,\alpha}) = \alpha$*. Then,* $\Pr(\chi^2_{n-1,1-\alpha/2} \le \frac{(n-1)s_n^2}{\sigma^2} \le \chi^2_{n-1,\alpha/2}) = 1 - \alpha$ *and, therefore,* $\Pr\left(\frac{(n-1)s_n^2}{\chi^2_{n-1,\alpha/2}} \le \sigma^2 \le \frac{(n-1)s_n^2}{\chi^2_{n-1,1-\alpha/2}}\right)$*. Now, when* $\alpha/2 = 0.025$*, chi-square table gives* $\chi^2_{24,0.025} = 39.364$ *and* $\chi^2_{24,0.975} = 12.401$ *so that the lower limit of the 95 percent CI is* $\frac{(n-1)s_n^2}{\chi^2_{n-1,0.025}} = \frac{24\times 2.36}{39.364} = 1.439$ *and the upper limit is* $\frac{(n-1)s_n^2}{\chi^2_{n-1,0.975}} = \frac{24\times 2.36}{12.401} = 4.567$*. Therefore, the 95 percent CI for the population variance is* $[1.439, 4.567]$ *and, because of the monotonic relationship between the variance and the standard deviation, the 95 percent CI for the population standard deviation is given by* $[\sqrt{1.439}, \sqrt{4.567}] = [1.200, 2.137]$*.*

To test the null hypothesis H_0, (a) find the distribution of "standardized" random variable $\frac{(n-1)s_n^2}{\sigma^2}$ when H_0 is true, i.e., $\sigma^2 = (2)^2 = 4$, (b) find the rejection region which is the region where the random variable $\frac{(n-1)s_n^2}{4}$ is unlikely (i.e., with the probability less than 5 percent) to fall into if H_0 is true, (c) look at the realized value of $\frac{(n-1)s_n^2}{4}$ and ask if the realized value of $\frac{(n-1)s_n^2}{4}$ is an unlikely value to happen if H_0 is true (by checking if $\frac{(n-1)s_n^2}{4}$ falls into the rejection region).

For (a), when H_0 is true, $\frac{(n-1)s_n^2}{4}$ is distributed according to the chi-square distribution with the degree of freedom equal to $n-1 = 24$. For (b), because $H_1 : \sigma > 2$, we consider one-sided test; namely, the very high value of $\frac{(n-1)s_n^2}{4}$ is considered to be evidence against H_0 but not the low value of $\frac{(n-1)s_n^2}{4}$. Under H_0, $\Pr\left(\frac{(n-1)s_n^2}{\sigma^2} \leq \chi^2_{n-1,\alpha}\right) = \Pr\left(\frac{24s_n^2}{4} \leq \chi^2_{24,\alpha}\right) = 1-\alpha$ for $\alpha = 0.05$, where $\chi^2_{24,0.05} = 36.415$. Therefore, the rejection region for $\frac{24s_n^2}{4}$ is given by $(36.415, \infty)$, i.e., we reject H_0 if $\frac{24s_n^2}{4} > 36.415$, or equivalently, $s_n^2 > 36.415/6 = 6.024$ because such a value of s_n^2 is unlikely to happen if H_0 is true. (c) The realized value of s_n^2 is 2.36, which does not fall in the rejection region (i.e., 2.36 belongs to the region which is not unlikely happen if H_0 is true) and hence there is not sufficient evidence against H_0. We do not reject H_0.