ISyE 6739 – Statistical Methods

Hypothesis Tests – One Population

Instructor: Kamran Paynabar
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

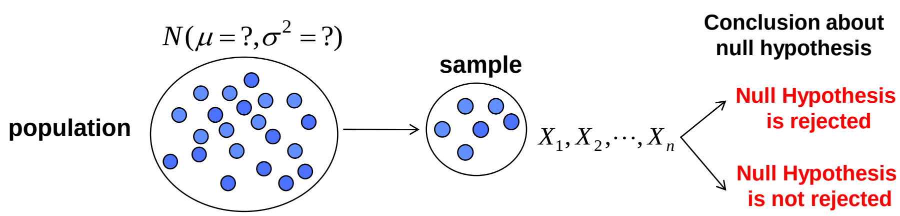
Kamran.paynabar@isye.gatech.edu
Office: Groseclose 436

Statistical Hypothesis Testing

- Statistical hypothesis testing of parameters are the fundamental methods used at the data analysis stage of a comparative experiment
- For example, comparing the mean of a population to a specified value

Null Hypothesis
Alternative Hypothesis

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{cases} \qquad \begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_a : \sigma^2 \neq \sigma_0^2 \end{cases}$$



A **statistical hypothesis** is a statement about the parameters of one or more populations.

Statistical Hypothesis Testing

Hypothesis-testing procedures rely on using the information in a random sample from the population of interest.

• If this information is *consistent* with the hypothesis, then we will conclude that the hypothesis is true; if this information is *inconsistent* with the hypothesis, we will conclude that the hypothesis is false.

$$H_0\colon \mu = 50 \text{ centimeters}$$
$$H_1\colon \mu \neq 50 \text{ centimeters}$$

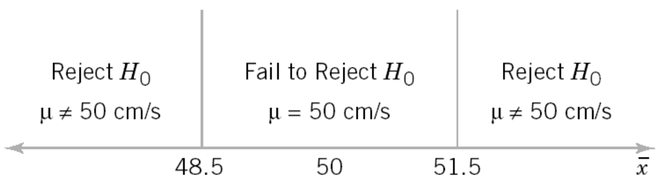


Figure 9-1 Decision criteria for testing H_0:μ = 50 centimeters per second versus H_1:$\mu \neq$ 50 centimeters per second.

Types of Hypotheses

Two-sided Alternative Hypothesis

X: inner-diameter of piston rings

$$H_0\colon \mu = 50$$ **null hypothesis**

$$H_1\colon \mu \neq 50$$ **alternative hypothesis**

One-sided Alternative Hypotheses

$$H_0: \mu = 10$$
$$H_1: \mu < 10$$

X: customers' waiting time in a bank

$$H_0: \mu = 100$$
$$H_1: \mu > 100$$

X: time to failure of machines

Hypothesis Testing Procedures

Design:

1) State the null and alternative hypotheses, and specify sample size n and significance level α.
2) Define the test statistic.
3) Find the distribution of the test statistic and the rejection region of H_0.

Example: Test on the mean of a Normal distribution with $\sigma^2 = 4$

$$H_0 : \mu = 50$$
$$H_1 : \mu \neq 50$$

$$n = 5$$
$$\alpha = 0.05$$

$$Z_0 = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Test Statistic

$$Z_0 \sim N(0,1)$$
$$|Z_0| > Z_{\alpha/2} = Z_{0.025} = 1.96$$

Rejection Region

Hypothesis Testing Procedures

Use:

4) Collect sample data and calculate the test statistic using the sample.
5) Compare the test statistic of the sample with the rejection region.
6) Make the decision and assess the risk.

Example (cont.): Test on the mean of a Normal distribution with $\sigma^2 = 4$

$$X_1 = 50.5$$
$$X_2 = 51.5$$
$$X_3 = 51$$
$$X_4 = 51.5$$
$$X_5 = 52$$

$$\rightarrow \bar{X} = 51.3 \rightarrow Z_0 = \frac{51.3 - 50}{2/\sqrt{5}} = 1.45$$

$$|Z_0| \not> Z_{\alpha/2} = 1.96$$

H0 is not rejected

Inference on the Mean of a Normal Population – Known Variance

Null Hypothesis	Test Statistic	Distribution under H0
$H_0: \mu = \mu_0$	$Z_0 = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$Z_0 \sim N(0,1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1: \mu \neq \mu_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$
$H_1: \mu > \mu_0$	$Z_0 > Z_{\alpha}$
$H_1: \mu < \mu_0$	$Z_0 < -Z_{\alpha}$

Critical values

Example

Aircrew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\overline{x} = 51.3$ centimeters per second. What conclusions should be drawn?

Example The response time of a distributed computer system is an important quality characteristic. The system manager wants to know whether the mean response time to a specific type of command exceeds 75 millisec. From past experience, he knows that the standard deviation of response time is 8 millisec.

If the command is executed 25 times and the response time for each trial is recorded. The sample average response time is 79.25 millisec. Formulate an appropriate hypothesis and test the hypothesis.

Hypothesis Testing Using Confidence Intervals Mean of a Normal Population – Known Variance

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0 \Rightarrow \mathrm{CI}: \left[\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

$$H_0: \mu = \mu_0 \quad H_1: \mu > \mu_0 \Rightarrow \mathrm{CI}: \left[\bar{x} - Z_{\alpha}\frac{\sigma}{\sqrt{n}}, +\infty\right)$$

Lower CI

$$H_0: \mu = \mu_0 \quad H_1: \mu < \mu_0 \Rightarrow \mathrm{CI}: \left(-\infty, \bar{x} + Z_{\alpha}\frac{\sigma}{\sqrt{n}}\right]$$

Upper CI

- **If the Confidence Interval does NOT include μ_0, then Reject H0**

Example The response time of a distributed computer system is an important quality characteristic. The system manager wants to know whether the mean response time to a specific type of command exceeds 75 millisec. From past experience, he knows that the standard deviation of response time is 8 millisec.

If the command is executed 25 times and the response time for each trial is recorded. The sample average response time is 79.25 millisec. Formulate an appropriate hypothesis and test the hypothesis.

Use CI method

Hypothesis Testing Using P-values

P-value definition:

The smallest significance level that would lead to the rejection of the null hypothesis

P-value calculation:

Pr(obtaining a test statistic at least as extreme as the one observed | H0 is true)

Conclusion based on P-value :

If the p-value < predefined α, **reject** the null hypothesis

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0$$

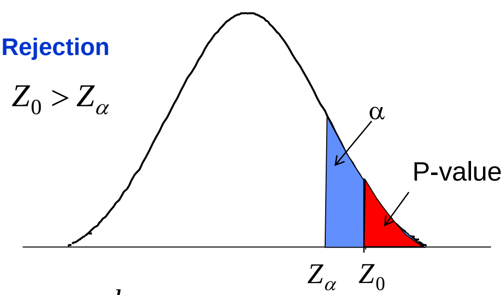
Observed statistic

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

Rejection

$$Z_0 > Z_\alpha$$

$$p - value = \Pr(Z > Z_0 | \mu = \mu_0)$$



$$Z_0 > Z_\alpha \Longleftrightarrow p - value < \alpha$$

P-Value for Inference on the Mean of Normal Distribution - Known Variance

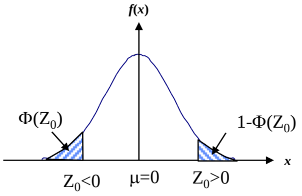
H_0: $\mu = \mu_0$

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

P-value=2[1-Φ(|Z_0|)] with two-sided H1: $\mu \neq \mu_0$

P-value=1-Φ(Z_0) for one-sided H1: $\mu > \mu_0$

P-value=Φ(Z_0) for one-sided H_1: $\mu < \mu_0$



- **A small p-value is a strong evidence against H0 while a large p-value means little evidence against H0.**

If p<α (predefined), we reject the null hypothesis.

Example

Aircrew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ and selects a random sample of $n = 25$ and obtains a sample average burning rate of $\bar{x} = 51.3$ centimeters per second. What conclusions should be drawn?

Use p-value method

Example The response time of a distributed computer system is an important quality characteristic. The system manager wants to know whether the mean response time to a specific type of command exceeds 75 millisec. From past experience, he knows that the standard deviation of response time is 8 millisec.

If the command is executed 25 times and the response time for each trial is recorded. The sample average response time is 79.25 millisec. Formulate an appropriate hypothesis and test the hypothesis.

Use p-value method

Inference on the Mean of a Normal Population – Known Variance

Null Hypothesis $H_0: \mu = \mu_0$

Test Statistic $Z_0 = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$

Distribution under H0 $Z_0 \sim N(0,1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)	Test using CI	P-values
$H_1: \mu \neq \mu_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$	$\left[\bar{x} - Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right]$	$2\left[1 - \Phi(\left\lvert\dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right\rvert)\right]$
$H_1: \mu > \mu_0$	$Z_0 > Z_\alpha$	$\left[\bar{x} - Z_\alpha \dfrac{\sigma}{\sqrt{n}}, +\infty\right)$	$1 - \Phi(\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}})$
$H_1: \mu < \mu_0$	$Z_0 < -Z_\alpha$	$\left(-\infty, \bar{x} + Z_\alpha \dfrac{\sigma}{\sqrt{n}}\right]$	$\Phi(\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}})$

Excel Function: F(x)=NORMDIST(z0,0,1,true)

Type I and Type II Errors

Conclusion

Reality	We Reject H0	We do NOT Reject H0
H0 is True	Type I error (α)	Confidence (1-α)
H0 is NOT True	Power (1-β)	Type II error (β)

Rejecting the null hypothesis H_0 when it is true is defined as a **type I error.**

Failing to reject the null hypothesis when it is false is defined as a **type II error.**

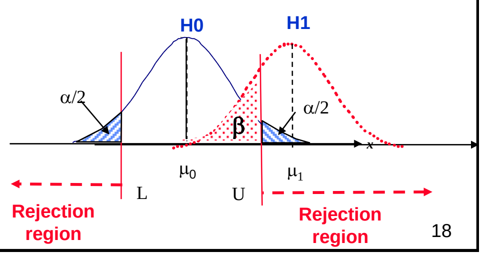
Two Types of Hypothesis Test Errors

- **Type I error (false reject/false alarm/false positive/significance level):**
  - α = P{type I error} = P{we reject H_0 |H_0 is true}
    = P{ test statistic falling in rejection region | H_0 is true}

- **Type II error (misdetection/false negative):**
  - β = P{type II error} = P{we fail to reject H_0 |H_0 is false}
    = P{ test statistic NOT falling in rejection region | H_0 is NOT true}

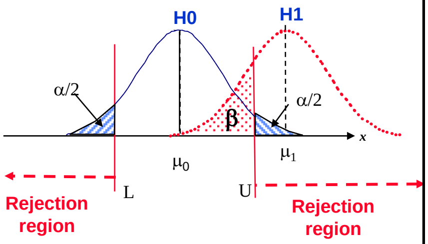
- **Power of the test (Correct detection):**
  - Power = 1- β = P{reject H0 | H_0 is NOT true}

Remember:
- α is determined by user
- The rejection limits (L & U) are based on H0 parameter!!



The Probability of Type II Error
Mean of a Normal Population- Known σ (Two-sided)

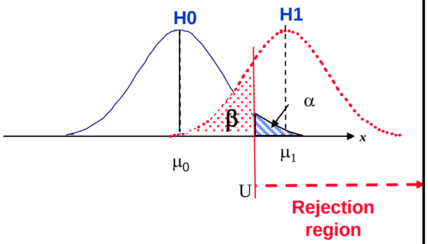
$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$
$$\downarrow$$
$$H_1 : \mu = \mu_1 = \mu_0 + \delta; (\delta \neq 0)$$



β = P{ test statistic NOT falling in rejection region | H_0 is NOT true}=....

The Probability of Type II Error
Mean of a Normal Population- Known σ (One-sided)

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu > \mu_0$$
$$\downarrow$$
$$H_1 : \mu = \mu_1 = \mu_0 + \delta; (\delta \neq 0)$$



β = P{ test statistic NOT falling in rejection region | H_0 is NOT true}=....

Example: The mean contents of coffee cans filled on a particular production line are being studied. Standards specify that the mean contents must be 16.0 oz, and from past experience it is known that the standard deviation of the can contents is 0.1 oz. The hypotheses are

H_0: $\mu=16.0$

H_1: $\mu\neq 16.0$

A random sample of nine cans is to be used, and type I error probability is specified as $\alpha=0.05$. What is the type II error if the true mean contents are $\mu_1=16.1$ oz or $\beta(16.1)$?

Properties of Type I & Type II Errors

- For a given sample size, one risk can only be reduced at the expense of increasing the other risk.
- For a given Type I error, the Type II error can be reduced by increasing the sample size at the price of increased inspection costs.

Choice of Sample Size

- Two-sided hypothesis test

$$n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2 \sigma^2}{\delta^2} \quad \text{where} \quad \delta = \mu - \mu_0 \qquad (9\text{-}19)$$

- One-sided hypothesis test

$$n = \frac{(z_{\alpha} + z_{\beta})^2 \sigma^2}{\delta^2} \quad \text{where} \quad \delta = \mu - \mu_0 \qquad (9\text{-}20)$$

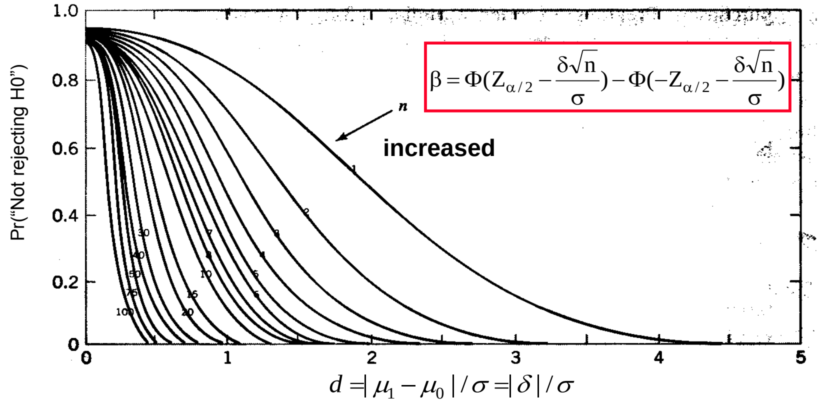
Example: Choice of Sample Size

Consider the rocket propellant problem of Example 9-2. Suppose that the analyst wishes to design the test so that if the true mean burning rate differs from 50 centimeters per second by as much as 1 centimeter per second, the test will detect this (i.e., reject H_0: $\mu = 50$) with a high probability, say 0.90. Now, we note that $\sigma = 2$, $\delta = 51 - 50 = 1$, $\alpha = 0.05$, and $\beta = 0.10$.

Operating Characteristic (OC) Curves

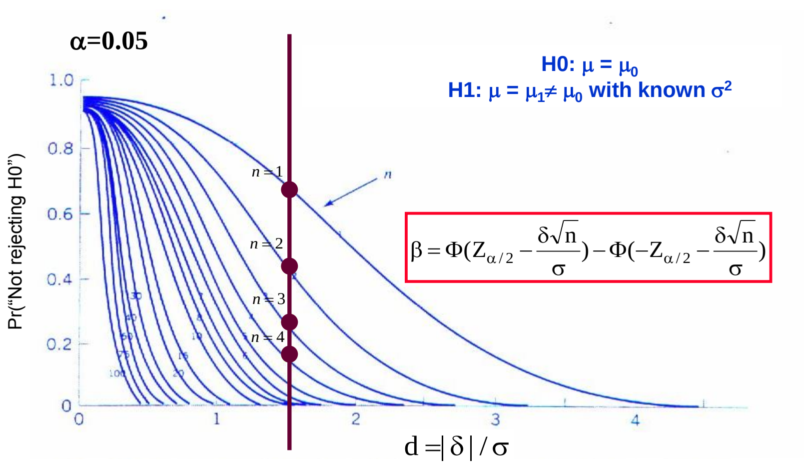
- Plot probability of "not rejecting H0" against δ
- The larger the mean shift, the smaller the type II error

H0: $\mu = \mu_0$
H1: $\mu = \mu_1 \neq \mu_0$ with known σ^2



Operating Characteristic (OC) Curves

- **The larger the sample size, the smaller the type II error**



Example: Suppose we wish to test the hypotheses

H_0: μ=15

H_1: μ≠15

where we know that σ^2=9.0. If the true mean is really 20, what sample size must be used to ensure that the probability of type II error is no greater than 0.10? Assume that α=0.05.

Inference on the Mean of a Normal Population – Unknown Variance

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \mu = \mu_0$	$t_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$t_0 \sim t(n-1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : \mu \neq \mu_0$	$\lvert t_0 \rvert > t_{\alpha/2,n-1}$
$H_1 : \mu > \mu_0$	$t_0 > t_{\alpha,n-1}$
$H_1 : \mu < \mu_0$	$t_0 < -t_{\alpha,n-1}$

Critical values

Hypothesis Testing Using Confidence Intervals Mean of a Normal Population – Unknown Variance

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0 \Rightarrow \text{CI}: \left[\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}\right]$$

$$H_0: \mu = \mu_0 \quad H_1: \mu > \mu_0 \Rightarrow \text{CI}: \left[\bar{x} - t_{\alpha,n-1}\frac{s}{\sqrt{n}}, +\infty\right)$$

Lower CI

$$H_0: \mu = \mu_0 \quad H_1: \mu < \mu_0 \Rightarrow \text{CI}: \left(-\infty, \bar{x} + t_{\alpha,n-1}\frac{s}{\sqrt{n}}\right]$$

Upper CI

- **If the Confidence Interval does NOT include μ_0, then Reject H0**

Example: The mean time it takes a crew to restart an aluminum rolling mill after a failure is of interest. The crew was observed over 25 occasions, and the results were $\bar{x}$ = 26.42 minutes and variance S^2 =12.28 minutes. If repair time is normally distributed,

- **Find a 95% confidence interval of the true mean repair time.**
- **Test the hypothesis that the true mean repair time is 30 minutes.**

Example

The increased availability of light materials with high strength has revolutionized the design and manufacture of golf clubs, particularly drivers. Clubs with hollow heads and very thin faces can result in much longer tee shots, especially for players of modest skills. This is due partly to the "spring-like effect" that the thin face imparts to the ball. Firing a golf ball at the head of the club and measuring the ratio of the outgoing velocity of the ball to the incoming velocity can quantify this spring-like effect. The ratio of velocities is called the coefficient of restitution of the club. An experiment was performed in which 15 drivers produced by a particular club maker were selected at random and their coefficients of restitution measured. In the experiment the golf balls were fired from an air cannon so that the incoming velocity and spin rate of the ball could be precisely controlled. It is of interest to determine if there is evidence (with $\alpha = 0.05$) to support a claim that the mean coefficient of restitution exceeds 0.82. The observations follow:

0.8411	0.8191	0.8182	0.8125	0.8750
0.8580	0.8532	0.8483	0.8276	0.7983
0.8042	0.8730	0.8282	0.8359	0.8660

The sample mean and sample standard deviation are $\bar{x} = 0.83725$ and $s = 0.02456$.

Inference on the Mean of a Normal Population – Unknown Variance

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \mu = \mu_0$	$t_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$t_0 \sim t(n-1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)	Test using CI	P-values
$H_1 : \mu \neq \mu_0$	$\lvert t_0 \rvert > t_{\alpha/2,n-1}$	$\left[\bar{x} - t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n}}\right]$	$2\left[1 - T_{n-1}\left(\left\lvert\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right\rvert\right)\right]$
$H_1 : \mu > \mu_0$	$t_0 > t_{\alpha,n-1}$	$\left[\bar{x} - t_{\alpha,n-1}\dfrac{s}{\sqrt{n}}, +\infty\right)$	$\left[1 - T_{n-1}\left(\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)\right]$
$H_1 : \mu < \mu_0$	$t_0 < -t_{\alpha,n-1}$	$\left(-\infty, \bar{x} + t_{\alpha,n-1}\dfrac{s}{\sqrt{n}}\right]$	$\left[T_{n-1}\left(\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)\right]$

CDF

$$T_{n-1}(a) = \int_{-\infty}^{a} f_{n-1}(t)\,dt$$

Inference on the Variance of a Normal Population

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \sigma^2 = \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 \sim \chi^2(n-1)$

Alternative Hypothesis	**Rejection/Critical Region** (H0 is rejected)
$H_1 : \sigma^2 \neq \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha/2,n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$
$H_1 : \sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha,n-1}$
$H_1 : \sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha,n-1}$

Critical values

Hypothesis Testing Using Confidence Intervals - Variance of a Normal Population

- Collect a sample and construct a 100(1- α)% CI

$$H_0 : \sigma^2 = \sigma_0^2,\quad H_1 : \sigma^2 \neq \sigma_0^2 \;\Rightarrow\; \text{CI} : \left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

$$H_0 : \sigma^2 = \sigma_0^2,\quad H_1 : \sigma^2 > \sigma_0^2 \;\Rightarrow\; \text{CI} : \left[\frac{(n-1)s^2}{\chi^2_{\alpha,n-1}}, +\infty\right)$$

Lower CI

$$H_0 : \sigma^2 = \sigma_0^2,\quad H_1 : \sigma^2 < \sigma_0^2 \;\Rightarrow\; \text{CI} : \left(0, \frac{(n-1)s^2}{\chi^2_{1-\alpha,n-1}}\right]$$

Upper CI

- If the Confidence Interval does NOT include σ_0^2, then Reject H0

Inference on the Variance of a Normal Population

Null Hypothesis	Test Statistic	Distribution under H0
$H_0 : \sigma^2 = \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 \sim \chi^2(n-1)$

Alternative Hypothesis	Rejection/Region (H0 is rejected)	Test using CI	CDF P-values
$H_1 : \sigma^2 \neq \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha/2, n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2, n-1}$	$\left[\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right]$	$2\left[1-\chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)\right]; if \chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right) > 0.5$ $2\left[\chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)\right]; if \chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right) < 0.5$
$H_1 : \sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi^2_{\alpha, n-1}$	$\left[\dfrac{(n-1)s^2}{\chi^2_{\alpha,n-1}}, +\infty\right)$	$\left[1-\chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)\right]$
$H_1 : \sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi^2_{1-\alpha, n-1}$	$\left(0, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha,n-1}}\right]$	$\left[\chi^2_{n-1}\left(\dfrac{(n-1)s^2}{\sigma_0^2}\right)\right]$

Example

An automatic filling machine is used to fill bottles with liquid detergent. A random sample of 20 bottles results in a sample variance of fill volume of $s^2 = 0.0153$ (fluid ounces)^2. If the variance of fill volume exceeds 0.01 (fluid ounces)^2, an unacceptable proportion of bottles will be underfilled or overfilled. Is there evidence in the sample data to suggest that the manufacturer has a problem with underfilled or overfilled bottles? Use $\alpha = 0.05$, and assume that fill volume has a normal distribution.

Inference on a Population Proportion

Null Hypothesis	Test Statistic	Asymptotic Distribution under H0
$H_0 : p = p_0$	$Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	$Z_0 \sim N(0,1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)
$H_1 : p \neq p_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$
$H_1 : p > p_0$	$Z_0 > Z_{\alpha}$
$H_1 : p < p_0$	$Z_0 < -Z_{\alpha}$

Critical values

Hypothesis Testing Using Confidence Intervals a Population Proportion

- **Collect a sample and construct a 100(1- α)% CI**

$$H_0 : p = p_0 \quad H_1 : p \neq p_0 \quad \Rightarrow \quad \text{CI}: \left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}, \hat{p} + Z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right]$$

$$H_0 : p = p_0 \quad H_1 : p > p_0 \quad \Rightarrow \quad \text{CI}: \left[\hat{p} - Z_{\alpha}\sqrt{\frac{p_0(1-p_0)}{n}}, +\infty\right)$$

Lower CI

$$H_0 : p = p_0 \quad H_1 : p < p_0 \quad \Rightarrow \quad \text{CI}: \left(-\infty, \hat{p} + Z_{\alpha}\sqrt{\frac{p_0(1-p_0)}{n}}\right]$$

Upper CI

- **If the Confidence Interval does NOT include p_0, then Reject H0**

One-sided 100(1-α)% CI for a Population Proportion with large sample size

$$X_1, X_2, \cdots, X_n \sim Bernoulli(p)$$

- Using CLT, we can estimate a CI for the population proportion.

$$\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N(p, \frac{p(1-p)}{n})$$

$$\hat{p} - Z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p$$

One-sided
Lower bound

$$p \le \hat{p} + Z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

One-sided
Upper bound

$$\hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Example

A semiconductor manufacturer produces controllers used in automobile engine applications. The customer requires that the process fallout or fraction defective at a critical manufacturing step not exceed 0.05 and that the manufacturer demonstrate process capability at this level of quality using $\alpha = 0.05$. The semiconductor manufacturer takes a random sample of 200 devices and finds that four of them are defective. Can the manufacturer demonstrate process capability for the customer?

Inference on a Population Proportion

Null Hypothesis

$H_0: p = p_0$

Test Statistic

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Asymptotic Distribution under H0

$Z_0 \sim N(0,1)$

Alternative Hypothesis	Rejection/Critical Region (H0 is rejected)	Test using CI	P-values
$H_1: p \neq p_0$	$\lvert Z_0 \rvert > Z_{\alpha/2}$	$\left[\hat{p} - Z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}, \hat{p} + Z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}\right]$	$2\left[1 - \Phi(\lvert Z_0 \rvert)\right]$
$H_1: p > p_0$	$Z_0 > Z_\alpha$	$\left[\hat{p} - Z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}, +\infty\right)$	$1 - \Phi(Z_0)$
$H_1: p < p_0$	$Z_0 < -Z_\alpha$	$\left(-\infty, \hat{p} + Z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}}\right]$	$\Phi(Z_0)$

Type II Error

For a two-sided alternative

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \quad (9\text{-}34)$$

If the alternative is $p < p_0$

$$\beta = 1 - \Phi\left(\frac{p_0 - p - z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \quad (9\text{-}35)$$

If the alternative is $p > p_0$

$$\beta = \Phi\left(\frac{p_0 - p + z_\alpha\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) \quad (9\text{-}36)$$

Choice of Sample Size

For a two-sided alternative

$$n = \left[\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p - p_0} \right]^2 \qquad (9\text{-}37)$$

For a one-sided alternative

$$n = \left[\frac{z_{\alpha}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p - p_0} \right]^2 \qquad (9\text{-}38)$$

Example

Consider the semiconductor manufacturer from Example 9-10. Suppose that its process fallout is really $p = 0.03$. What is the β-error for a test of process capability that uses $n = 200$ and $\alpha = 0.05$?

The β-error can be computed using Equation 9-35 as follows:

$$\beta = 1 - \Phi\left[\frac{0.05 - 0.03 - (1.645)\sqrt{0.05(0.95)/200}}{\sqrt{0.03(1-0.03)/200}} \right] = 1 - \Phi(-0.44) = 0.67$$

Thus, the probability is about 0.7 that the semiconductor manufacturer will fail to conclude that the process is capable if the true process fraction defective is $p = 0.03$ (3%). That is, the power of the test against this particular alternative is only about 0.3. This appears to be a large β-error (or small power), but the difference between $p = 0.05$ and $p = 0.03$ is fairly small, and the sample size $n = 200$ is not particularly large.

Example

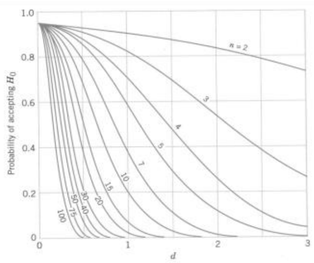
Suppose that the semiconductor manufacturer was willing to accept a β-error as large as 0.10 if the true value of the process fraction defective was $p = 0.03$. If the manufacturer continues to use $\alpha = 0.05$, what sample size would be required?

The required sample size can be computed from Equation 9-38 as follows:

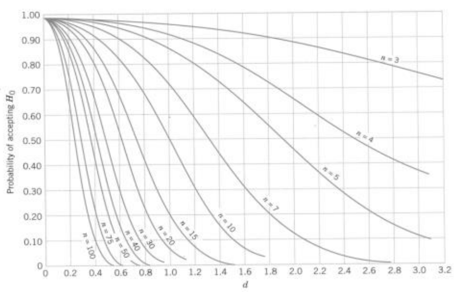
$$n = \left[\frac{1.645\sqrt{0.05(0.95)} + 1.28\sqrt{0.03(0.97)}}{0.03 - 0.05}\right]^2$$
$$\simeq 832$$

where we have used $p = 0.03$ in Equation 9-38. Note that $n = 832$ is a very large sample size. However, we are trying to detect a fairly small deviation from the null value $p_0 = 0.05$.

OC curve for t-test

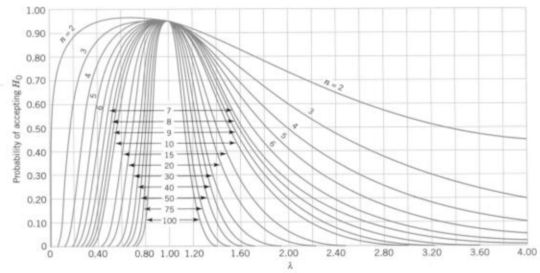


(e) O.C. curves for different values of n for the two-sided t-test for a level of significance $\alpha = 0.05$.

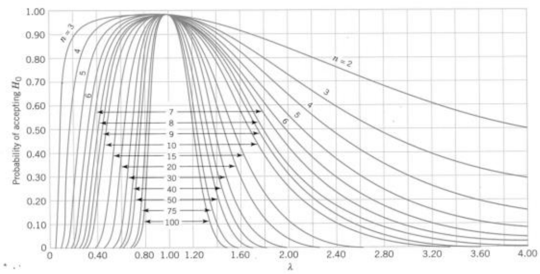


(f) O.C. curves for different values of n for the two-sided t-test for a level of significance $\alpha = 0.01$.

OC curve for Chi2-test



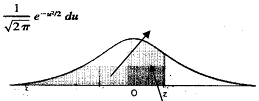
(i) O.C. curves for different values of n for the two-sided chi-square test for a level of significance $\alpha = 0.05$.



(j) O.C. curves for different values of n for the two-sided chi-square test for a level of significance $\alpha = 0.01$.

Appendix II Cumulative Standard Normal Distribution

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$



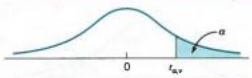
z	0.00	0.01	0.02	0.03	0.04	z
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.0
0.1	0.53983	0.54379	0.54776	0.55172	0.55567	0.1
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.2
0.3	0.61791	0.62172	0.62551	0.62930	0.63307	0.3
0.4	0.65542	0.65910	0.62276	0.66640	0.67003	0.4
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.5
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.6
0.7	0.75803	0.76115	0.76424	0.76730	0.77035	0.7
0.8	0.78814	0.79103	0.79389	0.79673	0.79954	0.8
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.9
1.0	0.84134	0.84375	0.84613	0.84849	0.85083	1.0
1.1	0.86433	0.86650	0.86864	0.87076	0.87285	1.1
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	1.2
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	1.3
1.4	0.91924	0.92073	0.92219	0.92364	0.92506	1.4
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	1.5
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	1.6
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	1.7
1.8	0.96407	0.96485	0.96562	0.96637	0.96711	1.8
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	1.9
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	2.0
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	2.1
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	2.2
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	2.3
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	2.4
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	2.5
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	2.6
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	2.7
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	2.8
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	2.9
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	3.0
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	3.1
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	3.2
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	3.3
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	3.4
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	3.5
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	3.6
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	3.7
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	3.8
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	3.9

Appendix II (*Continued*)

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

z	0.05	0.06	0.07	0.08	0.09	z
0.0	0.51994	0.52392	0.52790	0.53188	0.53586	0.0
0.1	0.55962	0.56356	0.56749	0.57142	0.57534	0.1
0.2	0.59871	0.60257	0.60642	0.61026	0.61409	0.2
0.3	0.63683	0.64058	0.64431	0.64803	0.65173	0.3
0.4	0.67364	0.67724	0.68082	0.68438	0.68793	0.4
0.5	0.70884	0.71226	0.71566	0.71904	0.72240	0.5
0.6	0.74215	0.74537	0.74857	0.75175	0.75490	0.6
0.7	0.77337	0.77637	0.77935	0.78230	0.78523	0.7
0.8	0.80234	0.80510	0.80785	0.81057	0.81327	0.8
0.9	0.82894	0.83147	0.83397	0.83646	0.83891	0.9
1.0	0.85314	0.85543	0.85769	0.85993	0.86214	1.0
1.1	0.87493	0.87697	0.87900	0.88100	0.88297	1.1
1.2	0.89435	0.89616	0.89796	0.89973	0.90147	1.2
1.3	0.91149	0.91308	0.91465	0.91621	0.91773	1.3
1.4	0.92647	0.92785	0.92922	0.93056	0.93189	1.4
1.5	0.93943	0.94062	0.94179	0.94295	0.94408	1.5
1.6	0.95053	0.95154	0.95254	0.95352	0.95448	1.6
1.7	0.95994	0.96080	0.96164	0.96246	0.96327	1.7
1.8	0.96784	0.96856	0.96926	0.96995	0.97062	1.8
1.9	0.97441	0.97500	0.97558	0.97615	0.97670	1.9
2.0	0.97982	0.98030	0.98077	0.98124	0.98169	2.0
2.1	0.98422	0.98461	0.98500	0.98537	0.98574	2.1
2.2	0.98778	0.98809	0.98840	0.98870	0.98899	2.2
2.3	0.99061	0.99086	0.99111	0.99134	0.99158	2.3
2.4	0.99286	0.99305	0.99324	0.99343	0.99361	2.4
2.5	0.99461	0.99477	0.99492	0.99506	0.99520	2.5
2.6	0.99598	0.99609	0.99621	0.99632	0.99643	2.6
2.7	0.99702	0.99711	0.99720	0.99728	0.99736	2.7
2.8	0.99781	0.99788	0.99795	0.99801	0.99807	2.8
2.9	0.99841	0.99846	0.99851	0.99856	0.99861	2.9
3.0	0.99886	0.99889	0.99893	0.99897	0.99900	3.0
3.1	0.99918	0.99921	0.99924	0.99926	0.99929	3.1
3.2	0.99942	0.99944	0.99946	0.99948	0.99950	3.2
3.3	0.99960	0.99961	0.99962	0.99964	0.99965	3.3
3.4	0.99972	0.99973	0.99974	0.99975	0.99976	3.4
3.5	0.99981	0.99981	0.99982	0.99983	0.99983	3.5
3.6	0.99987	0.99987	0.99988	0.99988	0.99989	3.6
3.7	0.99991	0.99992	0.99992	0.99992	0.99992	3.7
3.8	0.99994	0.99994	0.99995	0.99995	0.99995	3.8
3.9	0.99996	0.99996	0.99996	0.99997	0.99997	3.9

Percentage Points of the t Distribution[a]

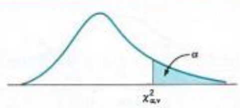


ν	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.49	4.019	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.20	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.992
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646

Chi-square Table

■ APPENDIX III

Percentage Points of the χ^2 Distribution[a]



ν	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.010	0.005
1	0.00 +	0.00 +	0.00 +	0.00 +	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.38	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.884	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93

Example

Suppose that the number of open circuits on a semiconductor wafer follow a Poisson distribution. A sample of 500 wafers indicate a total of 1038 open circuits. Test the hypothesis that the mean number of open circuits per wafer exceeds 2.0. (α = 0.05)

$$\begin{cases} H_0 : \lambda = 2 \\ H_1 : \lambda > 2 \end{cases}$$

$$MLE : \hat{\lambda} = \overline{X} \longrightarrow \begin{array}{l} E(\overline{X}) = E(X) = \lambda \\ \mathrm{var}(\overline{X}) = \mathrm{var}(X)/n = \lambda/n \end{array}$$

The ML estimators follow normal distribution when sample size is large. Therefore, under H0 we have

$$\overline{X} \sim N(2, 2/500) \longrightarrow Z_0 = \frac{\overline{X} - \lambda_0}{\sqrt{\lambda_0/n}} = \frac{\frac{1038}{500} - 2}{\sqrt{2/500}} = 1.2 < Z_{0.05} = 1.645$$

We cannot reject H0