Unit 5: Inferential statistics

- Confidence interval (CI)
  - CI on the population mean
  - CI on the population proportion
  - CI on the variance of a normal population
  - CI for the ratios of two variance

- Statistical hypothesis
  - Null and alternate hypothesis
  - The test statistic
  - Type I and type I errors
  - Level of significance and critical region
  - Hypothesis testing on the mean
  - Hypothesis testing on a proportion
  - Hypothesis testing on the variance
  - Goodness-of-fit

Confidence interval

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population $f(x)$ with mean μ and variance σ^2.

Recall that a point estimate is a single value estimate for a population parameter.

The most unbiased point estimate of the population mean μ is the sample mean $\bar{x}$

- A point estimate for a given sample is not exactly equal to the population parameter.
- We determine an interval within which we would expect to find the value of the parameter. Such an interval is called interval estimate.
- The interval estimate for a population parameter is called a confidence interval.

Below we construct confidence intervals on

1. The mean of a normal distribution, using either the normal distribution or the t distribution method.
2. The variance and standard deviation of a normal distribution.
3. A population proportion.

This confidence intervals play vital role in performing hypothesis testing.

Confidence interval (CI) on the population mean μ

A confidence interval estimate for μ is an interval of the form

$$l \leq \mu \leq u$$

where l and u are called lower- and upper-confidence limits, respectively.

- l and u are computed from the sample data.
- Because different samples will produce different l and u, these end-points are values of the random variable L and U, respectively.
- Suppose the probability

  $$P(L \leq \mu \leq U) = 1-\alpha \quad \text{is true,}$$

  where $0 \leq \alpha \leq 1$. There is a probability of $1-\alpha$ of selecting a sample for which the CI will contain the true value of μ.
- The interval $l \leq \mu \leq u$ is called a $100(1-\alpha)\%$ CI on μ.
- $1-\alpha$ is called the confidence coefficient or the degree of confidence.

Case i: Condifence interval on μ and, variance σ^2 known.

If $\bar{x}$ is the mean of a random sample of size n from a population $f(x)$ ($f(x)$ is normal or n is sufficiently large $(n \geq 30)$) with known variance σ^2, then

a $100(1-\alpha)\%$ Confidence interval for μ is given by

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

where $z_{\alpha/2}$ is the z-value leaving an area of $\alpha/2$ to the right.

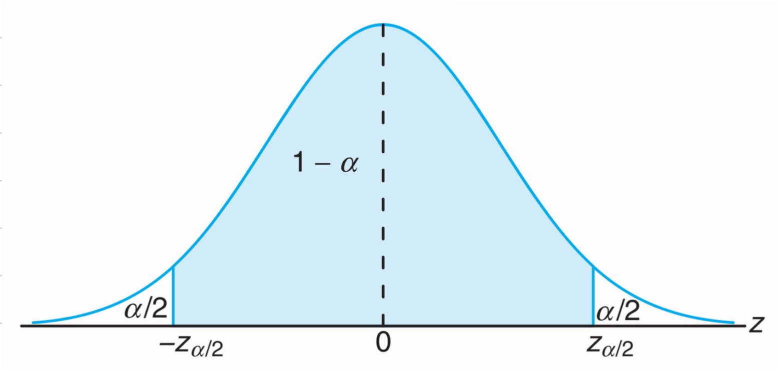
Pf: Let the sample mean $\bar{x}$ be the point estimate of μ.
We find e such that $\mu \in (\bar{x}-e, \bar{x}+e)$.

By CLT, the sampling distribution of $\bar{X}$ is approximately normally distributed with mean $\mu_{\bar{X}} = \mu$ and variance $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.
We may standardize $\bar{X}$ by

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

The r.v. Z has a standard normal distribution.

Write $z_{\alpha/2}$ for the z-value above which we find an area of $\alpha/2$ under the normal curve,



We see from the figure that

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$$

$$\Rightarrow P\left(-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Thus, a $100(1-\alpha)\%$ CI on μ is

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Here lower confidence limit $l = \bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n}$ and

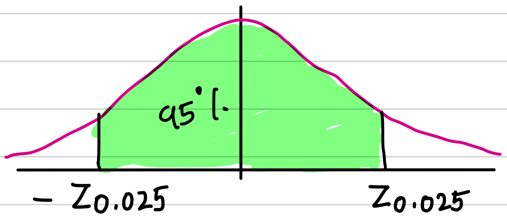
upper confidence limit $u = \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n}$.

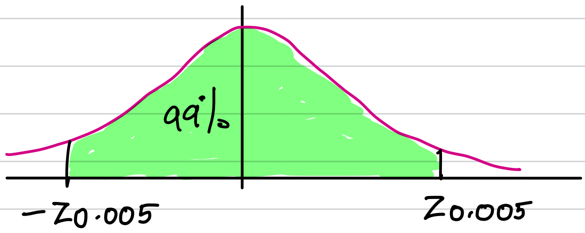
In particular,

If $\alpha = 0.05$, we have 95% confidence interval, and when $\alpha = 0.01$, we have 99% confidence interval.

From the cumulative distribution table

$z_{0.05/2} = z_{0.025} = 1.96$ and $z_{0.01/2} = z_{0.005} = 2.575 \approx 2.58$





95% confidence interval :	$\bar{x} - 1.96\,\sigma/\sqrt{n} < \mu < \bar{x} + 1.96\,\sigma/\sqrt{n}$
99% confidence interval :	$\bar{x} - 2.58\,\sigma/\sqrt{n} < \mu < \bar{x} + 2.58\,\sigma/\sqrt{n}$

Maximum error of estimate e on μ, σ^2 known

Given a $100(1-\alpha)\%$ confidence interval, the maximum error of estimate

$$e = |\bar{x} - \mu| \text{ will not exceed } z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

i.e $\quad e \leq z_{\alpha/2}\,\sigma/\sqrt{n}$

Sample size for specified error on μ, σ^2 known

For a point estimate $\bar{x}$, we can be $100(1-\alpha)\%$ confident that the error will not exceed specified amount e when the sample size is

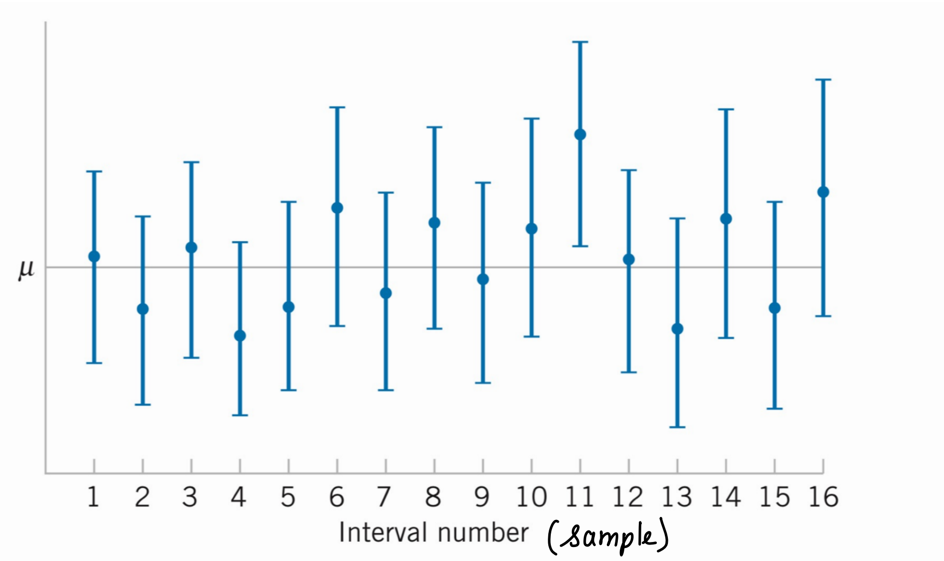
$$n = \left(\frac{z_{\alpha/2}\,\sigma}{e}\right)^2.$$

Interpretating a confidence interval

A CI is a random interval because the endpoints of the interval are random variables.

If several samples are collected and a $100(1-\alpha)\%$ CI for μ is computed for each sample, then $100(1-\alpha)\%$ of these invervals contain the true value of μ.

Figure shows the interval estimate of μ for different values of $\bar{x}$.



- dot at each interval indicates position of $\bar{x}$ for that random sample
- Width of each interval is same, depends on the choice of $z_{\alpha/2}$
- For a selection of $z_{\alpha/2}$, $100(1-\alpha)\%$ of the invervals will cover μ.

Ex1: The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per milliliter. Find the 95% and 99% confidence intervals for the mean zinc concentration in the river. Assume that the population standard deviation is 0.3 gram per milliliter.

i) How large a sample is required if we want to be 95% confident that our estimate of μ is off by less than 0.05?

Soln: Sample size $= 36$, population S.D $\sigma = 0.3$ gm/ml the point estimate of μ is $\bar{x} = 2.6$ gm/ml

The 95% confidence interval is

$$\bar{x} - z_{0.05/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{0.05/2}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow 2.6 - (1.96)\frac{0.3}{\sqrt{36}} < \mu < 2.6 + (1.96)\frac{0.3}{\sqrt{36}}$$

$$\Rightarrow 2.5 < \mu < 2.7$$

The 99% confidence interval is

$$\bar{x} - z_{0.01/2}\,\sigma/\sqrt{n} < \mu < \bar{x} + z_{0.01/2}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow 2.6 - (2.575)\frac{0.3}{\sqrt{36}} < \mu < 2.6 + (2.575)\frac{0.3}{\sqrt{36}}$$

$$\Rightarrow 2.47 < \mu < 2.73$$

i) Given error $|\bar{x} - \mu| < 0.05$

$\therefore$ Sample size, $n = \left(\frac{z_{0.05/2}\,\sigma}{e}\right)^2 = \left(\frac{(1.96)(0.3)}{0.05}\right)^2 \approx 139$

One-sided confidence bounds on μ, σ^2 known

Above what we discussed is a two-sided CI. It is also possible to obtain one-sided confidence bounds for μ by setting either $l = -\infty$ or $u = -\infty$ and replacing $z_{\alpha/2}$ by z_α.

A $100(1-\alpha)\%$ upper-confidence bound for μ is

$$\mu \le \bar{x} + z_\alpha\,\sigma/\sqrt{n}$$

and a $100(1-\alpha)\%$ lower-confidence bound for μ is

$$\bar{x} - z_\alpha\,\sigma/\sqrt{n} \le \mu$$

Case ii : Large sample confidence interval for μ, σ^2 unknown

$(n \geq 40)$

When n is large, we can replace σ by the sample standard deviation S. Thus

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has an approximate standard normal distribution. Consequently

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

is large sample $100(1-\alpha)\%$ confidence interval for μ.

Large sample one-sided confidence bounds on μ, σ^2 unknown

A $100(1-\alpha)\%$ upper-confidence bound for μ is

$$\mu \leq \bar{x} + z_\alpha \, s/\sqrt{n}$$

and a $100(1-\alpha)\%$ lower-confidence bound for μ is

$$\bar{x} - z_\alpha \, s/\sqrt{n} \leq \mu$$

Maximum error of estimate e on μ, σ^2 unknown, n large

Given a $100(1-\alpha)\%$ confidence interval, the maximum error of estimate

$e = |\bar{x} - \mu|$ will not exceed $z_{\alpha/2} \frac{s}{\sqrt{n}}$

i.e $\quad e \leq z_{\alpha/2} \, s/\sqrt{n}$

Case iii : Confidence interval on μ of a normal distribution, σ^2 unknown (small sample size n)

When n is small, a different distribution must be employed to construct the CI.

t - distribution

Let $X_1, X_2, \cdots, X_n$ be a random sample from a normal distribution with unknown mean μ and unknown variance σ^2. The random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a t distribution with $n-1$ degree of freedom.

The t probability density function is

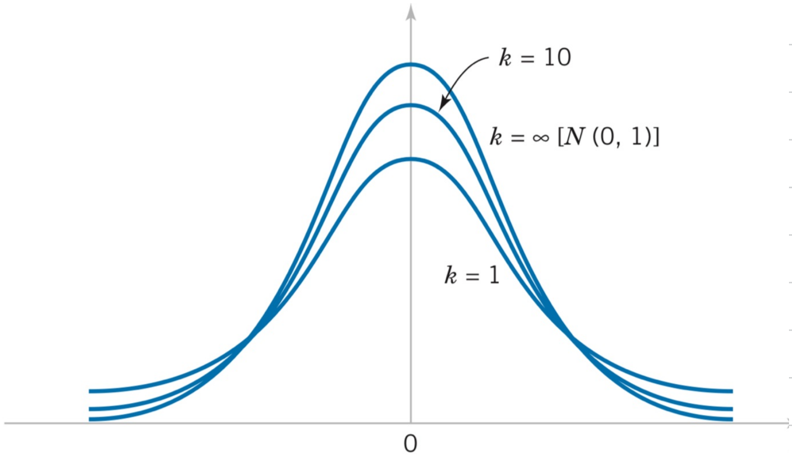
$$f(x) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k}\,\Gamma(k/2)} \cdot \frac{1}{[(x^2/k)+1]^{(k+1)/2}}, \qquad -\infty < x < \infty$$

where k is the number of degree of freedom. The mean and variance of t distribution are zero and $k/(k-2)$ (for $k>2$), respectively.

Properties of t - distribution

- It is symmetric and unimodal, maximum ordinate value is reached when $\mu = 0$.
- It is similar to the standard normal distribution. However, the t distribution has heavier tails than the normal.
- As the no. of degree of freedom $k \to \infty$, the t-distribution approaches standard normal distribution.

Probability density functions of several t distributions



Confidence interval on μ, σ^2 unknown (n is small)

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance σ^2, then

a $100(1-\alpha)\%$ confidence interval on μ is given by

$$\bar{x} - t_{\alpha/2, n-1}\, s/\sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2, n-1}\, s/\sqrt{n}$$

where $t_{\alpha/2, n-1}$ is the upper $100\alpha/2$ percentage point of t distribution with $n-1$ degree of freedom.

One-sided t confidence bounds

A $100(1-\alpha)\%$ upper-confidence bound for μ is

$$\mu \le \bar{x} + t_{\alpha, n-1}\, s/\sqrt{n}$$

and a $100(1-\alpha)\%$ lower-confidence bound for μ is

$$\bar{x} - t_{\alpha, n-1}\, s/\sqrt{n} \le \mu$$

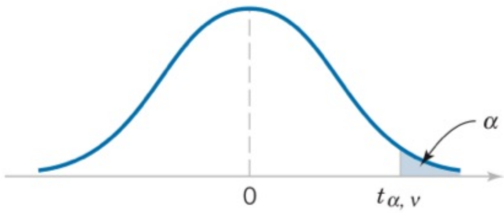


Table V Percentage Points $t_{\alpha,\nu}$ of the t Distribution

ν \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.

Ex2: The contents of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 liters. Find a 95% confidence interval for the mean contents of all such containers, assuming an approximately normal distribution.

Soln: Given: Seven samples 9.8, 10.2, 10.4, 9.8, 10, 10.2, 9.6

$$\bar{x} = \frac{9.8+10.2+10.4+9.8+10+10.2+9.6}{7} = 10$$

$$s^2 = \frac{(9.8-10)^2+(10.2-10)^2+(10.4-10)^2+(9.8-10)^2+(10-10)^2+(10.2-10)^2+(9.6-10)^2}{7-1}$$

$= 0.08$

$s = 0.283$

For above table $t_{0.025} = 2.447$ for $v = 6$.

Hence 95% confidence interval of μ is

$$\bar{x} - t_{0.025}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{0.025}\, s/\sqrt{n}$$

$$\Rightarrow \quad 10 - 2.447\left(\frac{0.283}{\sqrt{7}}\right) < \mu < 10 + 2.447\left(\frac{0.283}{\sqrt{7}}\right)$$

$$\Rightarrow \quad 9.74 < \mu < 10.26$$

Confidence interval for a population proportion p

Recall that the sampling distribution of the sample proportion $\hat{p}$ is approximately normal with mean p and variance $\frac{p(1-p)}{n}$, if p is not close to 0 or 1 or n is relatively large. Typically, np and $n(1-p) \geq 5$.

Thus, if n is large, the distribution of

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

is approximately standard normal.

Large sample confidence interval on $\hat{p}$

If $\hat{p}$ is the proportion of observation in a random sample of size n that belongs to a class of interest, then an approximate $100(1-\alpha)\%$ confidence interval on the proportion p of the population is

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

Maximum error of estimate e on p

Given a $100(1-\alpha)\%$ confidence interval, the maximum error of estimate

$e = |\bar{x} - \mu|$ will not exceed $z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

i.e. $e \leq z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Ex 3 : In a random sample of n = 500 families owning television sets in the city of Hamilton, Canada, it is found that x = 340 subscribe to HBO. Find a 95% confidence interval for the actual proportion of families with television sets in this city that subscribe to HBO.

Soln: $n = 500$, $\hat{p} = \frac{x}{n} = \frac{340}{500} = 0.68$.

The 95% confidence interval

$$\hat{p} - Z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + Z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\Rightarrow \quad 0.68 - 1.96\sqrt{\frac{0.68(1-0.68)}{500}} \leq p \leq 0.68 + 1.96\sqrt{\frac{0.68(1-0.68)}{500}}$$

$$\Rightarrow \quad 0.6391 \leq p \leq 0.7209.$$

Confidence interval on the variance of a normal distribution

Let the population be modelled by normal distribution.

χ^2 distribution

Let $X_1, X_2, \cdots X_n$ be a random sample from a normal distribution with mean μ and variance σ^2, let S^2 be the sample variance. Then the random variable

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma^2}$$

has a chi-square (χ^2) distribution with $n-1$ degree of freedom.

The probability density fn of a χ^2 r.v. is

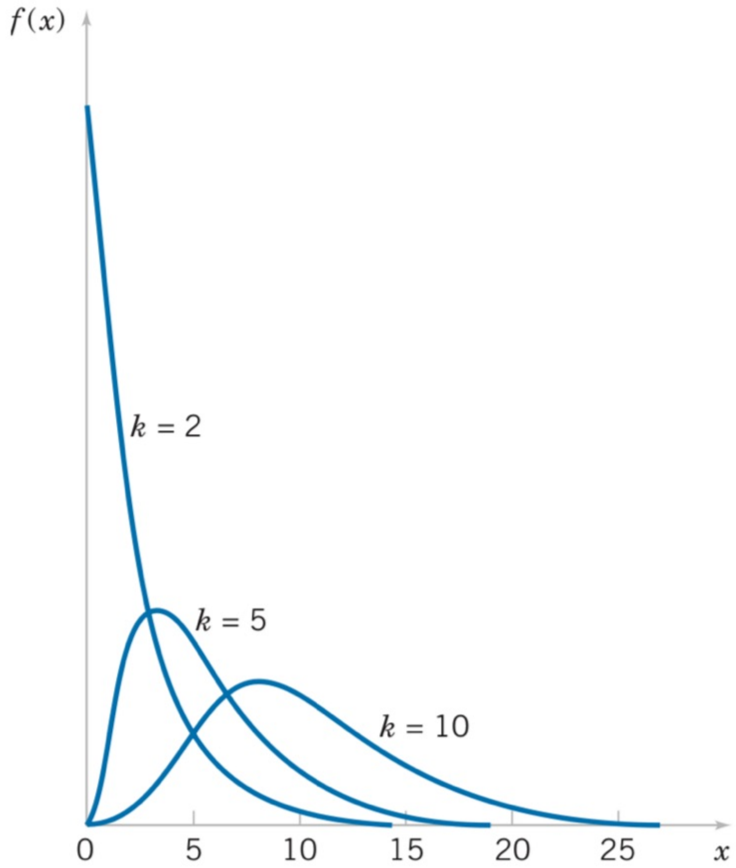
$$f(x) = \frac{1}{2^{k/2}\,\Gamma(k/2)}\, x^{(k/2)-1}\, e^{-x/2}, \quad x > 0$$

where k is the no. of degrees of freedom.

Properties of χ^2 distribution

- The mean $= k$
- Variance $= 2k$
- The r.v. is non-negative
- The probability distribution is skewed to the right.
- As k increases, the distribution becomes more symmetric
- As $k \to \infty$, the χ^2 distribution approaches normal distribution

Probability density functions of several χ^2 distributions



Confidence interval on the variance

If s^2 is the variance of a random sample of size n from a normal distribution, then

a $100(1-\alpha)\%$ confidence interval on σ^2 is

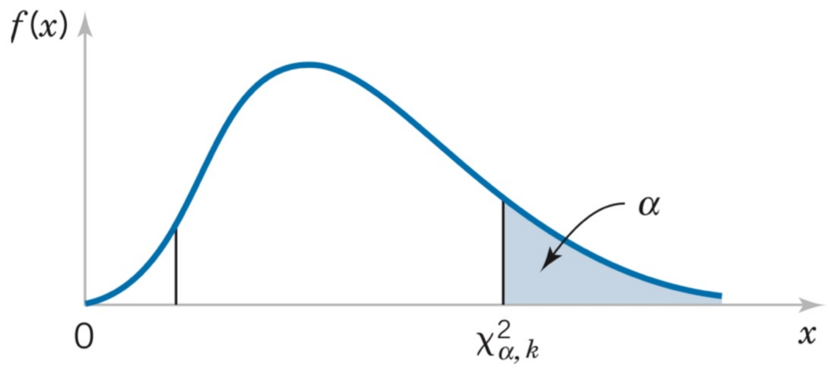
$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,n-1}}$$

where $\chi^2_{\alpha/2}$ and $\chi^2_{1-\alpha/2}$ are χ^2-values with $n-1$ degrees of freedom, leaving areas of $\alpha/2$ and $1-\alpha/2$, respectively, to the right.

An approximate $100(1-\alpha)\%$ CI for σ is

$$\frac{\sqrt{n-1}\,s}{\sqrt{\chi^2_{\alpha/2,\,n-1}}} \le \sigma \le \frac{\sqrt{n-1}\,s}{\sqrt{\chi^2_{1-\alpha/2,\,n-1}}}$$

percentage point $\chi^2_{\alpha,k}$



One-sided confidence bounds on the variance

The $100(1-\alpha)\%$ lower and upper confidence bounds on σ^2 are

$$\frac{(n-1)s^2}{\chi^2_{\alpha,n-1}} \le \sigma^2 \quad \text{and} \quad \frac{(n-1)s^2}{\chi^2_{1-\alpha,n-1}} \ge \sigma^2$$

respectively.

Ex: The following are the weights, in decagrams, of 10 packages of grass seed distributed by a certain company: 46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2, and 46.0. Find a 95% confidence interval for the variance of the weights of all such packages of grass seed distributed by this company, assuming a normal population.

Soln: Sample is

46.4, 46.1, 45.8, 47, 46.1, 45.9, 45.8, 46.9, 45.2, 46.0

$n=10$; $\bar{x} = \frac{\sum x_i}{n} = 46.12$; $s^2 = \frac{\sum (x_i-\bar{x})^2}{n-1} = 0.286$

95% confidence interval, ($\alpha = 0.05$)

$$\frac{9s^2}{\chi^2_{0.025,9}} \le \sigma^2 \le \frac{9s^2}{\chi^2_{1-0.025,9}},$$

Using table $\chi^2_{0.025,9} = 19.023$

$\chi^2_{0.975,9} = 2.700$

$\Rightarrow \quad 0.135 \le \sigma^2 \le 0.953$

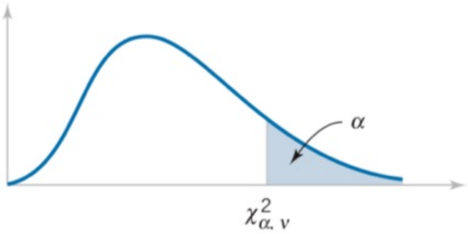


Table Percentage Points $\chi^2_{\alpha,\nu}$ of the Chi-Squared Distribution

ν \ α	.995	.990	.975	.950	.900	.500	.100	.050	.025	.010	.005
1	.00+	.00+	.00+	.00+	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.01	.02	.05	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.07	.11	.22	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.21	.30	.48	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.41	.55	.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.68	.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

ν = degrees of freedom.

Confidence interval for the ratios of two variance

Let σ_1^2 and σ_2^2 be the variances of two normal populations. Let s_1^2 and s_2^2 be the variances of independent samples of size n_1 and n_2, respectively, from normal populations. Then the estimate of the ratio $\frac{\sigma_1^2}{\sigma_2^2}$ is $\frac{s_1^2}{s_2^2}$. Hence the statistic $\frac{S_1^2}{S_2^2}$ is called an estimator of $\frac{\sigma_1^2}{\sigma_2^2}$

The interval estimate of $\frac{\sigma_1^2}{\sigma_2^2}$ is given by the statistic

$$F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

The r.v. F has F-distribution with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom.

A $100(1-\alpha)\%$ confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$ is

$$\frac{s_1^2}{s_2^2} \frac{1}{f_{\alpha/2, v_1, v_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} \frac{1}{f_{1-\alpha/2, v_1, v_2}}$$

where $f_{\alpha/2, v_1, v_2}$ and $f_{1-\alpha/2, v_1, v_2}$ are f-values with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom, leaving areas $\alpha/2$ and $1-\alpha/2$, respectively, to the right.

Note that $$f_{\alpha/2, v_2, v_1} = \frac{1}{f_{1-\alpha/2, v_1, v_2}}$$

Statistical hypothesis

A statistical hypothesis is a statement concerning about the parameters of one or more population.

Ex: 1) $H: p = 0.1$, where p is population proportion of defective items.

2) $H: \mu = 1100$ hr, where μ is average life span of electric bulbs manufactured by a company.

3) Let μ_1 and μ_2 denote the true average tensile strengths of two different materials.
hypothesis might be assertion that $\mu_1 - \mu_2 \geq 2$, or another statement is $\mu_1 - \mu_2 = 0$.

4) Let vaccine A be 25% effective on a certain viral infection after 2 yrs. To determine new and more expensive vaccine B is superior in providing protection against the same virus after 2yrs.
hypothesis might be assertion that $P(B) > 0.25$, or $P(B) = 0.25$.

Steps involved in hypothesis testing

1) Formulate the hypothesis to be tested.
2) Determine the appropriate test statistic and calculate it using the sample data
3) Compare test statistic to critical region to draw initial conclusion.
4) Calculate p-value

5) Conclusion, written in terms of the original problem.

1. Null and Alternate hypothesis

In any hypothesis-testing problem, there are always two competing hypothesis under consideration:

a) The null hypothesis (H_0)

b) The alternate hypothesis (H_1)

The alternate hypothesis H_1 usually represents the question to be answered or the theory to be tested.

The null hypothesis H_0 nullifies or opposes H_1 and is often the logical complement to H_1.

Basically, null hypothesis is tested for possible rejection under the assumption that it is true.

Ex : 1) If a person encounter a jury trial, the

H_0 : defendent is innocent

H_1 : defendent is guilty

2) A factory has a machine that dispenses 100ml of juice in a bottle. The employ believes that the average amount of fluid is not 100ml.

$$H_0 : \mu = 100\,ml$$

$$H_1 : \mu \neq 100\,ml$$

3) A company manufactures car batteries with an average life span of 2 or more yrs. An engineer believes this value to be less.

$H_0 : \mu \geq 2$ yrs (we consider $H_0 : \mu = 2$ yrs)

$H_1 : \mu < 2$ yrs

Note: Null hypothesis H_0 will often be stated with equality sign.

2. The test statistic

A test statistic assesses how consistent your sample data with the null hypothesis in a hypothesis test.

The test statistic is a function of sample statistic, its value which calculated using the sample data tell us by how much sample diverges from the null hypothesis.

As a test statistic value become more extreme, it indicates large differences between your sample data and the null hypothesis.

If the difference is significant, the null hypothesis is rejected. If it is not, then the null hypothesis is not rejected.

Ex 1) A factory has a machine that dispenses 80ml of juice in a bottle. The employ believes that the average amount of fluid is not 80 ml.

$H_0 : \mu = 100$ml

$H_1 : \mu \neq 100$ml

Using 40 samples, he measures the average amount dispensed by the machine to be 78 ml with SD 2.5ml

Given: $\mu = 80$ml, $\bar{x} = 78$ ml, $s = 2.5$ml and $n = 40$

Since n is large, the test statistic is Z-distribution and the value, $z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{78 - 80}{2.5/\sqrt{40}} = -5.06$

tells us by how much sample mean diverges from the null hypothesis.

Errors in hypothesis testing.

When we perform hypothesis test, we make one of two decisions:

a) reject the null hypothesis or

b) fail to reject the null hypothesis.

Because our decision is based on the sample rather than the entire population, there is a possibility we may make one of the following errors.

1) Type 1 error :

The error occurs if we reject null hypothesis, when it is true.

2) Type 2 error :

This error occurs if we donot reject null hypothesis, when it is false.

Thus, in testing any statistical hypothesis, four different situations determine whether the final decision is correct or error. These situations are

Decision	H_0 is true	H_0 is false
Fail to reject H_0	Correct decision	type II error
Reject H_0	type I error	Correct decision

Level of significance or α-error and β-error

The probability of type I error is called the level of significance or α-error.

$$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$$

The probability of type II error is called β-error

$$\beta = P(\text{type II error}) = P(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false})$$

Ex 2) Suppose $f(x)$ is a normal population with mean 50 and standard deviation 2.5. Test the hypothesis that $\mu = 50$ against the alternative that $\mu \neq 50$ if the size of the random sample is 10.

Given: Null hypothesis, $H_0 : \mu = 50$; $\sigma = 2.5$ and $n = 10$

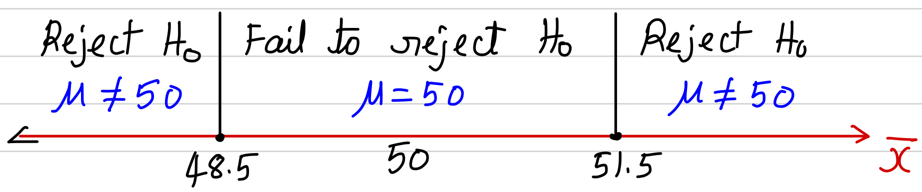
Alternate hypothesis $H_1 : \mu \neq 50$

The sample mean can take on many different values. Suppose that

if $48.5 \leq \bar{x} \leq 51.5$, we will not reject $H_0 : \mu = 50$, and

if $\bar{x} < 48.5$ or $\bar{x} > 51.5$ we will reject H_0 in favour of $H_1 : \mu \neq 50$

This is illustrated as below:



The level of significance or α-level:

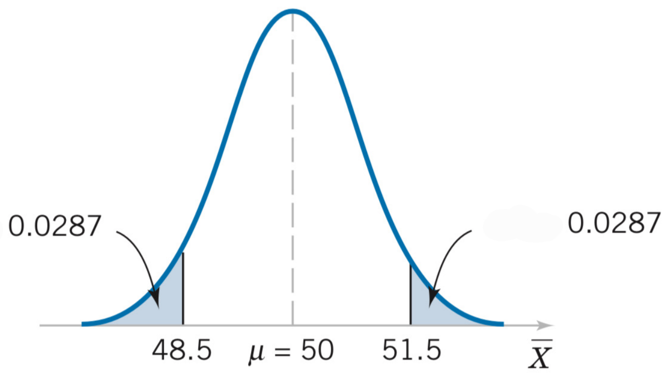
$$\begin{aligned}
\alpha &= P(\text{reject } H_0 \text{ when } \mu = 50) \\
&= P(\bar{x} < 48.5 \text{ or } \bar{x} > 51.5 \text{ when } \mu = 50) \\
&= P(\bar{x} < 48.5 \text{ when } \mu = 50) + P(\bar{x} > 51.5 \text{ when } \mu = 50)
\end{aligned}$$

Z-values are

$$Z_1 = \frac{48.5 - \mu}{\sigma/\sqrt{n}} = \frac{48.5 - 50}{2.5/\sqrt{10}} = -1.9$$

and $$Z_2 = \frac{51.5 - \mu}{\sigma/\sqrt{n}} = \frac{51.5 - 50}{2.5/\sqrt{10}} = 1.9$$

$$\therefore \quad \alpha = P(Z < -1.9) + P(Z > 1.9) = 0.0287 + 0.0287$$
$$= 0.0574$$



Note that α decreasing with increase in n.

The β-level: We can't calculate β-level unless we have a specific alternate hypothesis; that is, we must have a particular value of μ.

Suppose that if $\mu = 52$ and $48.5 \le \bar{x} \le 51.5$, then we donot reject H_0. This leads to type II error.

$$\therefore \quad \beta = P(\text{fail to reject } H_0 \text{ when } \mu = 52)$$
$$= P(48.5 \le \overline{X} \le 51.5 \text{ when } \mu = 52)$$

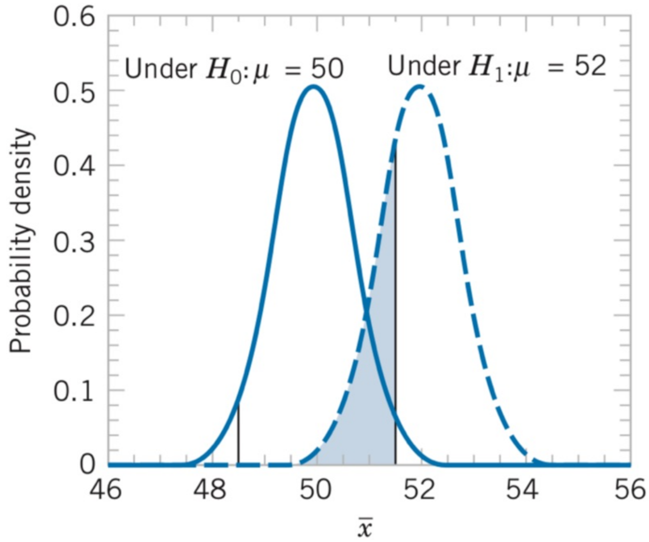
The Z-values are

$$Z_1 = \frac{48.5 - \mu}{\sigma/\sqrt{n}} = \frac{48.5 - 52}{2.5/\sqrt{10}} = -4.43$$

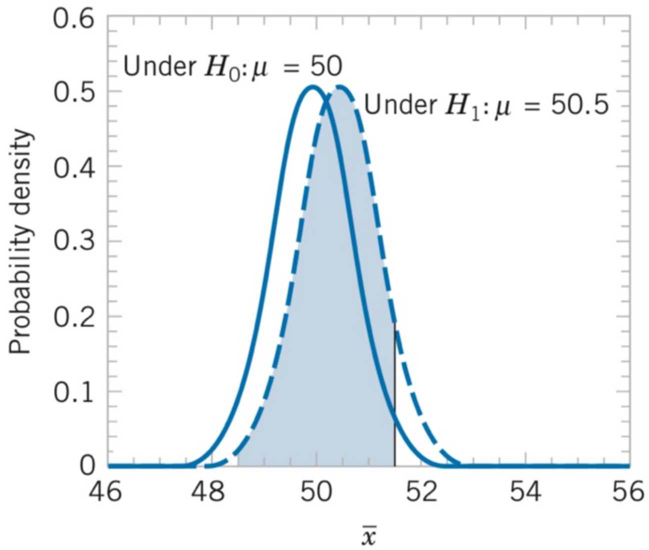
$$Z_2 = \frac{51.5 - 52}{2.5/\sqrt{10}} = -0.63$$

$$\therefore \beta = P(-4.43 \le Z \le -0.63) = \phi(-0.63) - \phi(-4.43)$$
$$= 0.2643$$

By symmetry, if the true value of the mean is $\mu = 48$, the value of β will also be 0.2643.



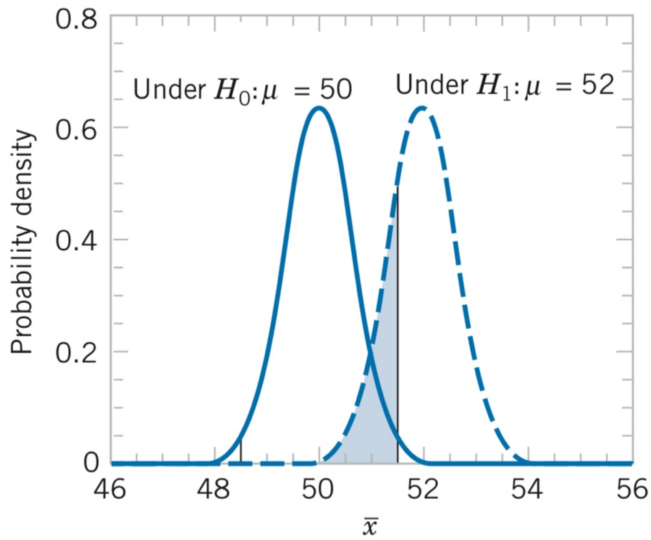
The probability of type II error when $\mu = 52$ and $n = 10$



The probability of type II error when $\mu = 50.5$ and $n = 10$

The probability of making type II error β increases rapidly as the true value of μ approaches the hypothesized value.

For Ex: $\beta = P(48.5 \le \bar{X} \le 51.5 \text{ when } \mu = 50.5) = 0.8923$.



The probability of type II error when $\mu = 52$ and $n = 16$.

Note: 1) Type I and type II errors are related. A decrease in the probability of one generally results in an increase in the probability of other.

2) An increase in sample size reduces β, provided that α is held constant.

Critical region and critical value

Critical region :

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected.

i.e, if the observed test statistic is in the critical region that we reject the null hypothesis and accept the alternative hypothesis.

Acceptance region :

It is a set of values for the test statistic for which the null hypothesis is not reject.

Critical value:

The value of test statistic which separates the critical region and the acceptance region is called the critical value.

In the above example 2,

the acceptance region is $48.5 \le \bar{X} \le 51.5$

Z-values are $-1.9 \le Z \le 1.9$

the critical region is $\bar{X} > 51.5$ or $\bar{X} < 48.5$

Z-values are $Z > 1.9$ or $Z < -1.9$

the critical values are -1.9 and 1.9

One – and Two–Tailed Tests

- A test of any statistical hypothesis where the alternate is one sided, such as

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta > \theta_0$$

or perhaps

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta < \theta_0$$

is called a one-tailed test.

- A test of any statistical hypothesis where the alternative is two-sided, such as

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

is called a two-tailed test.

Note:

1) The critical region for the alternative hypothesis $\theta > \theta_0$ lies in the right tail of the distribution of the test statistic.

2) The critical region for the alternative hypothesis $\theta < \theta_0$ lies in the left tail

3) The critical region for the alternative hypothesis $\theta \neq \theta_0$ lies in both tails.

Ex 1: A large manufacturing firm is being charged with discrimination in its hiring practices.

(a) What hypothesis is being tested if a jury commits a type I error by finding the firm guilty?

(b) What hypothesis is being tested if a jury commits a type II error by finding the firm guilty?

Soln: a) Null hypothesis, H_0: The firm is not guilty

Alternate hypothesis, H_1: The firm is guilty

H_0 is tested.

b) Null hypothesis, H_0: The firm is guilty

Alternate hypothesis, H_1: The firm is not guilty

H_0 is tested.

Ex 2: A fabric manufacturer believes that the proportion of orders for raw material arriving late is p = 0.6. If a random sample of 10 orders shows that 3 or fewer arrived late, the hypothesis that p = 0.6 should be rejected in favor of the alternative p < 0.6. Use the binomial distribution.

(a) Find the probability of committing a type I error if the true proportion is p = 0.6.

(b) Find the probability of committing a type II error for the alternatives p = 0.3.

Soln: Null hypothesis, $p = 0.6$

Alternate hypothesis, $p < 0.6$

Let X be the no. of orders arrived late.

Sample size, $n = 10$

We reject the null hypothesis when $X \leq 3$

a) $P(\text{Type I error}) = P(\text{reject } H_0 \text{ when } p = 0.6)$

$= P(X \leq 3 \text{ when } p = 0.6)$

$= \sum_{x=0}^{3} b(x; n, p)$

$$= b(0;10,0.6) + b(1;10,0.6) + b(2;10,0.6) + b(3;10,0.6)$$

$$= (0.4)^{10} + 10(0.6)(0.4)^9 + \binom{10}{2}(0.6)^2(0.4)^8 + \binom{10}{3}(0.6)^3(0.4)^7$$

$$= 0.05476$$

b) $P(\text{Type II error}) = P(\text{fail to reject } H_0, \text{ when } p=0.3)$

$$= P(X>3, \text{ when } p=0.3)$$

$$= \sum_{x=4}^{10} b(x;n,p)$$

$$= 1 - \sum_{x=0}^{3} b(x;10,0.3)$$

$$= 1 - b(0;10,0.3) - b(1;10,0.3) - b(2;10,0.3) - b(3;10,0.3)$$

$$= 1 - (0.7)^{10} - 10(0.3)(0.7)^9 - \binom{10}{2}(0.3)^2(0.7)^8 - \binom{10}{3}(0.3)^3(0.7)^7$$

$$= 1 - 0.6496$$

$$= 0.3504$$

Preselection of a significance level and critical region

We can control the maximum risk of making type I error by preselecting a significance level.

Let us fix the significance level as α, we say α level of significance. Then $100(1-\alpha)\%$ confidence interval for the parameter θ constitute the acceptance region for the null hypothesis $\theta = \theta_0$.

That is,

Let $\hat{\theta}$ be the value of a test statistic in the hypothesis testing. Then:

i) The value $\hat{\theta}$ is said to be in the acceptance region for the α level of sigificance if it falls within the $100(1-\alpha)\%$ confidence interval.

ii) The value $\hat{\theta}$ is said to be in the critical region for the α level of sigificance if it falls outside the $100(1-\alpha)\%$ the confidence interval.

Note: The widely used procedure in hypothesis testing is to use significance level $\alpha = 0.05$. This value has evolved through experience, and may not be appropriate for all situations.

P-values in hypothesis tests

The fixed significance level is nice, but it is inadequate to conclude whether the computed value of the test statistic is just barely in the rejection region or whether is very far into this region.

This approach may be unsatisfactory because some decision

makers might be uncomfortable with the risk implied by $\alpha = 0.05$.
To avoid this P-value approach is adopted.

The P-value is the probability that the test statistic take on a value that is at least as extreme as the observed value of the statistic when H_0 is true.

Defn: The P-value is the smallest level of significance that would lead to rejection of the null hypothesis H_0 with the given data.

Ex: Suppose $H_0: \mu = 50$, $H_1: \mu \neq 50$, $\bar{x} = 48.5$ and $n = 16$. $\sigma = 2.5$

Test statistic

$$z = \frac{48.5 - 50}{2.5/\sqrt{16}} = -2.40$$

$$\text{P-value} = 2P(Z \leq -2.40) = 0.0164$$

P-value (area shaded)

1. Testing hypotheses on the mean, σ^2 known (Z tests)

Let us consider that the population is normal or sample size n is sufficiently large ($n \geq 30$).

Null hypothesis : $H_0 : \mu = \mu_0$

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

Let $\bar{x}$ be the value of the sample mean. The corresponding z-value is:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

For α level of significance

Alternate hypothesis	$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$
Tests	Two-tailed test	lower-tailed test	upper-tailed test
Critical region	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$	$z_0 < -z_\alpha$	$z_0 > z_\alpha$
Acceptance region	$-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}$	$z_0 \geq -z_\alpha$	$z_0 \leq z_\alpha$
P-value	probability above $\lvert z_0 \rvert$ and probability below $-\lvert z_0 \rvert$, $P = 2(1 - \phi(\lvert z_0 \rvert))$	probability below z_0, $P = \phi(z_0)$	probability above z_0, $P = 1 - \phi(z_0)$

2. Testing hypotheses on the mean, σ^2 Unknown (Z tests)

Let us consider that the sample size n is large ($n \geq 40$)

Null hypothesis : $H_0 : \mu = \mu_0$

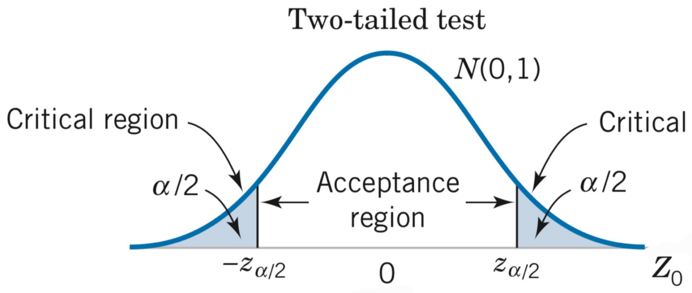
Test statistic : $Z_0 = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$, where S is sample variance

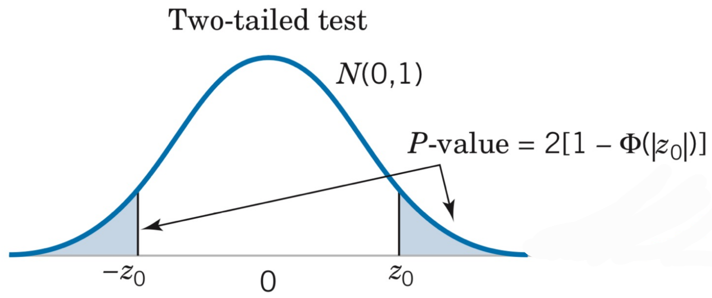
Rest is same as the previous case.

Critical regions and P-values (graphical representation)

Hypothesis, $H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$





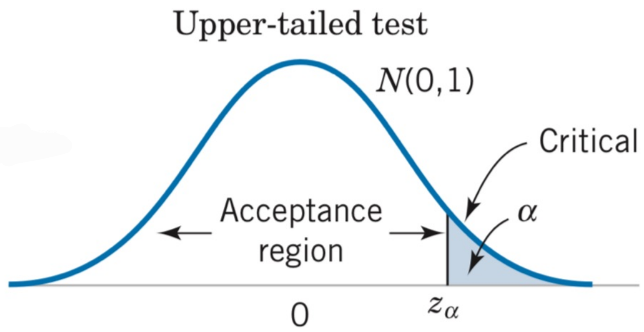
Hypothesis

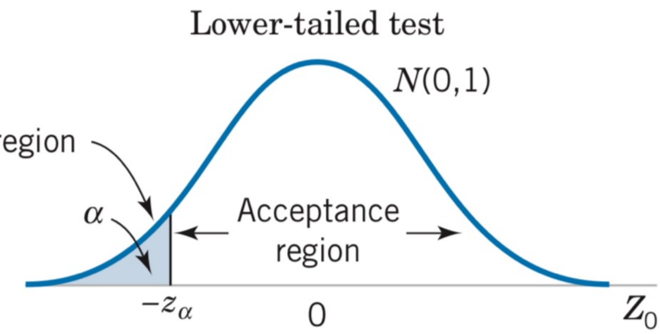
$H_0 : \mu = \mu_0$

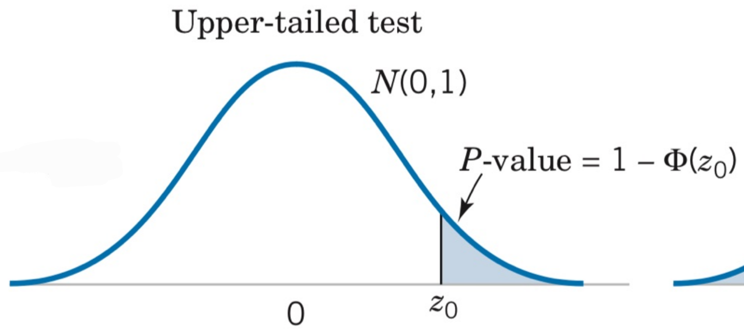
$H_1 : \mu > \mu_0$

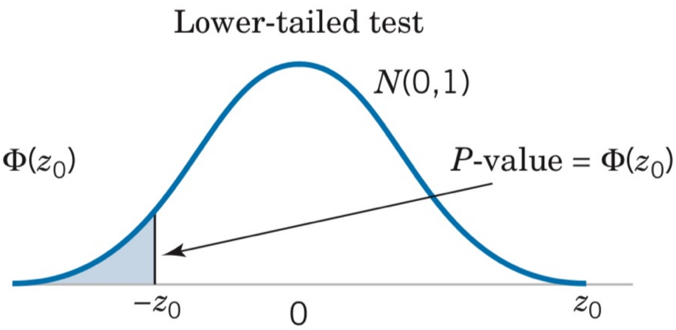
$H_0 : \mu = \mu_0$

$H_1 : \mu < \mu_0$









3. Testing hypotheses on the mean, σ^2 unknown (t-test)

Let us consider that the population is at least approximately normal.

Null hypothesis : $H_0 : \mu = \mu_0$

Test statistic : $T_0 = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$, where S is sample variance

Let $\bar{x}$ be the value of the sample mean, let s be its variance.
The corresponding t-value is:

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

For α level of significance

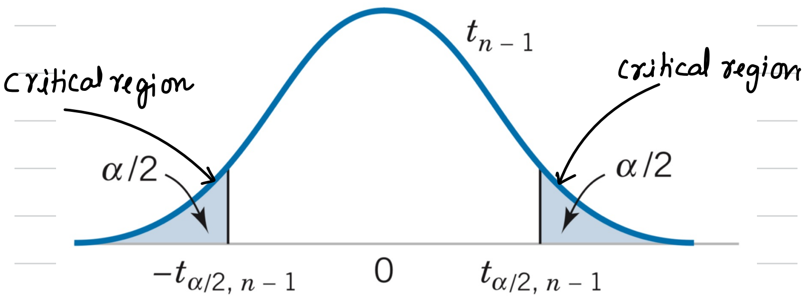
Alternate hypothesis	$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$
Tests	Two-tailed test	lower-tailed test	upper-tailed test
Critical region	$t_0 > t_{\alpha/2, n-1}$ or $t_0 < -t_{\alpha/2, n-1}$	$t_0 < -t_{\alpha, n-1}$	$t_0 > t_{\alpha, n-1}$
Acceptance region	$-t_{\alpha/2, n-1} \leq t_0 \leq t_{\alpha/2, n-1}$	$t_0 \geq -t_{\alpha, n-1}$	$t_0 \leq t_{\alpha, n-1}$
P-value	probability above $\lvert t_0 \rvert$ and probability below $-\lvert t_0 \rvert$,	probability below t_0	probability above t_0,

Critical regions and P-values (graphical representation)

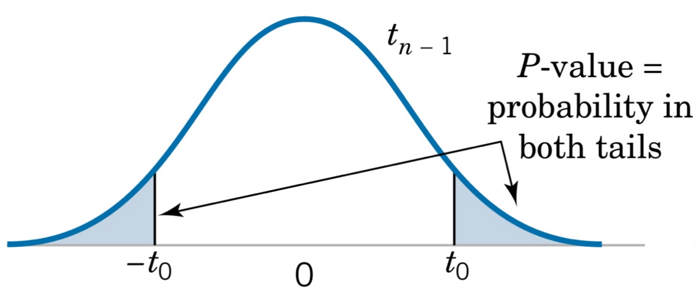
Hypothesis, $H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$

two-tailed test



Two-tailed test



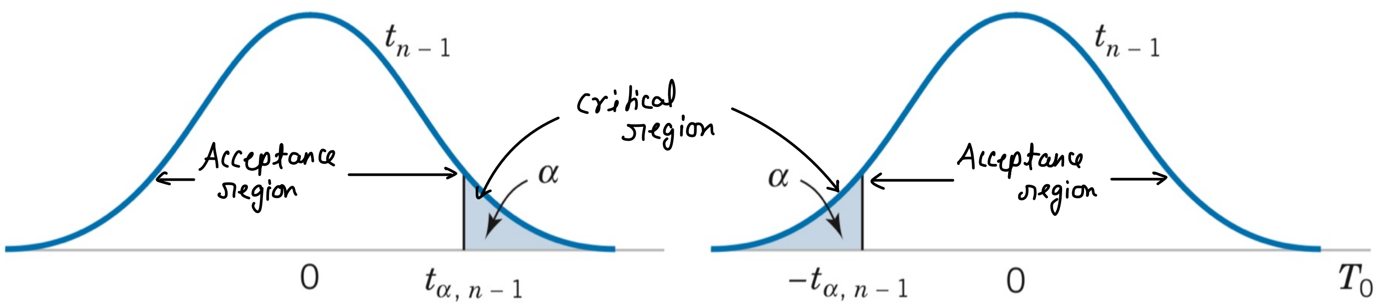
Hypothesis

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

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

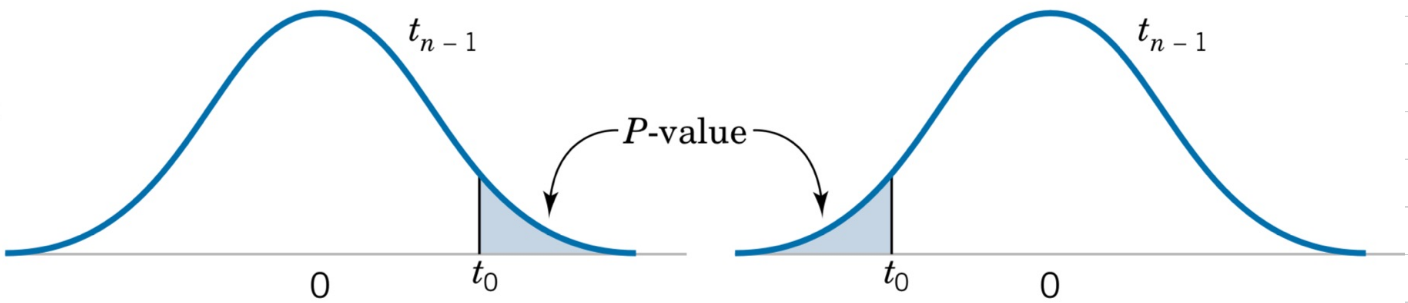
Upper tailed test

Lower tailed test



One-tailed test

One-tailed test



4. Testing hypotheses on a proportion, large sample size

Null hypothesis : $H_0 : p = p_0$

Let x be no. of observations in a random sample of size n that belongs to a class of interest. Then

Sample proportion $\hat{p} = \frac{x}{n}$, its distribution is approximately normal as n is large.

$\therefore$ Test statistic : $Z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

For α level of significance

Alternate hypothesis	$H_1 : p \neq p_0$	$H_1 : p < p_0$	$H_1 : p > p_0$
Tests	Two-tailed test	lower-tailed test	upper-tailed test
Critical region	$Z_0 > Z_{\alpha/2}$ or $Z_0 < -Z_{\alpha/2}$	$Z_0 < -Z_\alpha$	$Z_0 > Z_\alpha$
Acceptance region	$-Z_{\alpha/2} \leq Z_0 \leq Z_{\alpha/2}$	$Z_0 \geq -Z_\alpha$	$Z_0 \leq Z_\alpha$
P-value	Probability above $\lvert Z_0 \rvert$ and probability below $-\lvert Z_0 \rvert$, $P = 2(1 - \phi(\lvert Z_0 \rvert))$	probability below Z_0 $P = \phi(Z_0)$	Probability above Z_0, $P = 1 - \phi(Z_0)$

Small sample tests on a binomial proportion

Tests on a proportion when the sample size n is small are based on the binomial distribution, not the normal approximation to the binomial.

5. Testing hypotheses on the variance (χ^2 test)

We wish to test the hypothesis that the variance of a normal population σ^2 equals a specified value, say σ_0^2 or equivalently $\sigma = \sigma_0$.

Null hypothesis : $H_0 : \sigma^2 = \sigma_0^2$

Test statistic : $\chi_0^2 = \dfrac{(n-1)\,S^2}{\sigma_0^2}$

Let s be the value of the sample variance. Then the χ^2-value is

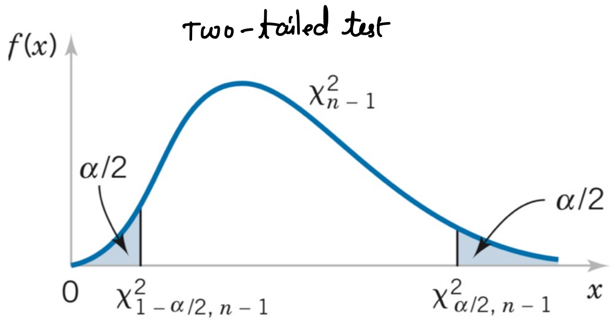
$$\chi_0^2 = \frac{(n-1)\,s^2}{\sigma_0^2}$$

For α level of significance

Alternate hypothesis	$H_1 : \sigma^2 \neq \sigma_0^2$	$H_1 : \sigma^2 < \sigma_0^2$	$H_1 : \sigma^2 > \sigma_0^2$
Tests	Two-tailed test	lower-tailed test	upper-tailed test
Critical region	$\chi_0^2 > \chi^2_{\alpha/2,\,n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2,\,n-1}$	$\chi_0^2 < \chi^2_{1-\alpha,\,n-1}$	$\chi_0^2 > \chi^2_{\alpha,\,n-1}$
Acceptance region	$\chi^2_{1-\alpha/2,\,n-1} \leq \chi_0^2 \leq \chi^2_{\alpha/2,\,n-1}$	$\chi_0^2 \geq \chi^2_{1-\alpha,\,n-1}$	$\chi_0^2 \leq \chi^2_{\alpha,\,n-1}$

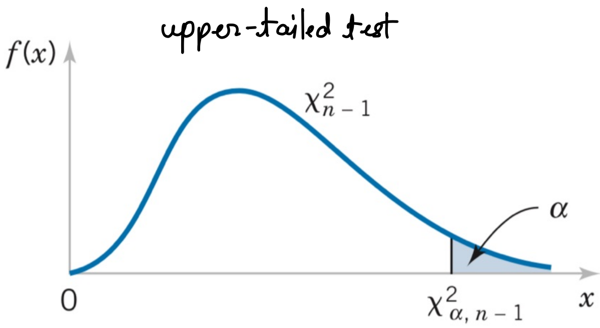
Critical region (graphical representation)

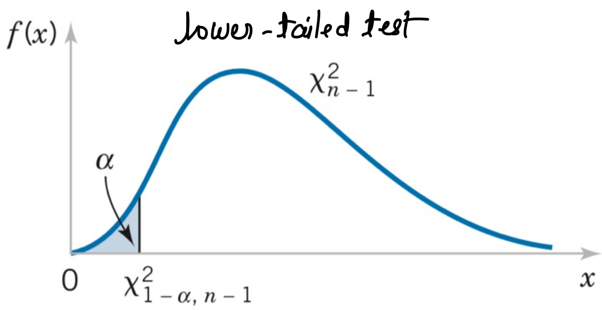
Null hypothesis, $H_0 : \sigma^2 = \sigma_0^2$
Alternate hypothesis, $H_1 : \sigma^2 \neq \sigma_0^2$



$H_0: \sigma^2 = \sigma_0^2$
$H_1: \sigma^2 > \sigma_0^2$

$H_0 : \sigma^2 = \sigma_0^2$
$H_1: \sigma^2 < \sigma_0^2$





Problems

Ex 1: A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5 kilogram. Test the hypothesis that μ = 8 kilograms against the alternative that μ= 8 kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms. Use a 0.01 level of significance.

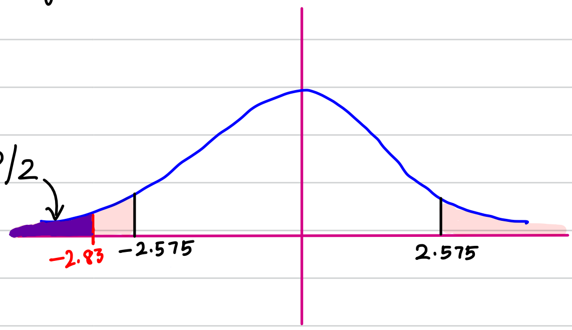
Soln: Null hypothesis, $H_0 : \mu = 8$ kg

Alternate hypothesis, $H_1 : \mu \neq 8$ kg

Given: $\sigma = 0.5$ kg,

Sample size, $n = 50$,

Sample mean, $\bar{x} = 7.8$ kg



Level of significance: $\alpha = 0.01$.

Test: Two-tailed test (z-test)

Test statistic: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{7.8 - 8}{0.5/\sqrt{50}} = -2.83$ (since n is large)

Critical region: $z < -z_{0.01/2}$ and $z > z_{0.01/2}$

i.e $z < -2.575$ and $z > 2.575$

Decision: Since $z = -2.83 < -2.575$, the value is in the critical region. ∴ we reject H_0 and conclude that the average breaking strength is not equal to 8, but in fact less than 8.

P-value: $P = P(|Z| > 2.83) = 2P(Z < -2.83)$

$= 2(1 - \Phi(|-2.83|)) = 0.0046$

Which allow us to reject $H_0 : \mu = 8$ kg at a level of significance smaller than 0.01.

Ex2: A random sample of 100 recorded deaths in the United States during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

Soln: Null hypothesis, $H_0 : \mu = 70$ years

Alternate hypothesis, $H_1 : \mu > 70$ years

Given : $\sigma = 8.9$ years

sample size, $n = 100$

sample mean, $\bar{x} = 71.8$ years

Level of significance : $\alpha = 0.05$

Test : Upper-tailed test (z-test)

Test statistic : $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$ (since n is large)

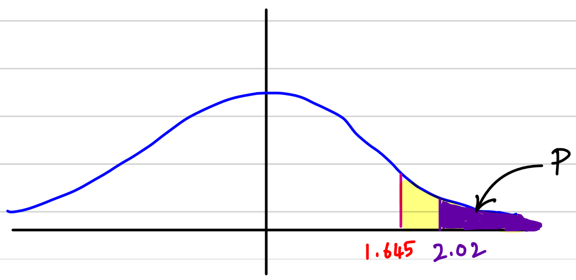
Critical region : $z > z_{0.05}$

i.e $z > 1.645$

Decision : Since the test value $z = 2.02 > 1.645$, we reject null hypothesis H_0 and conclude that the mean life span today is greater than 70 years

P-value : $P = P(z > 2.02) = 1 - \phi(2.02) = 0.0217$

This implies that the evidence is in favor of H_1 is even stronger than suggested by a 0.05 level of significance.



Ex 3: The Edison Electric Institute has published figures on the number of kilowatt hours used annually by various home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

Soln: Null hypothesis, $H_0 : \mu = 46$ kilowatt hour

Alternate hypothesis, $H_1 : \mu < 46$ kilowatt hour

Given : sample size, $n = 12$

sample mean, $\bar{x} = 42$ kilowatt hour

sample std deviation, $s = 11.9$ kilowatt hour

Level of significance : $\alpha = 0.05$

Test : Lower-tailed test (t-test)

Test statistic : $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ with 11 degree of freedom

(Since σ is unknown, populatn is normal)

Computation : $= \dfrac{42 - 46}{11.9/\sqrt{12}} = -1.16$

Critical region : $t < -t_{\alpha, n-1}$

i.e $t < -t_{0.05, 11}$

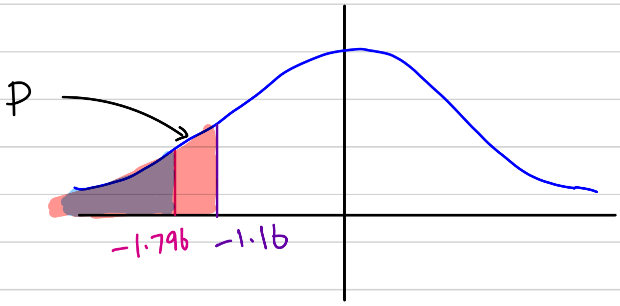
$\Rightarrow t < -1.796$

P-value : $P = P(T < -1.16)$

≈ 0.135

(This value can't be found using the given table, I have given it for reference)

Decision : Since the value $t = -1.16 > -1.796$, we donot reject null hypothesis H_0 and conclude that the average no. of kilowatt hours used annually by home vaccum cleaner is not significantly less than 46.



Ex4: A commonly prescribed drug for relieving nervous tension is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance.

Soln: Null hypothesis, $H_0 : p = 0.6$

Alternate hypothesis, $H_1 : p > 0.6$

Given: sample size, $n = 100$ adults

no. of individual received relief, $x = 70$

sample proportion, $\hat{p} = \frac{x}{n} = 0.7$

Level of significance: $\alpha = 0.05$

Test: Upper-tailed test (z-test)

Test statistic: $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{p(1-p)}{n}}}$ (Since n is large)

Computation: $= \dfrac{0.7 - 0.6}{\sqrt{\dfrac{0.6(1-0.6)}{100}}} = 2.04$

Critical region: $Z > Z_{0.05}$

i.e $Z > 1.645$

P-value: $P = P(z > 2.04)$
< 0.0207

Decision: As the value $z = 2.04 > 1.645$, we reject H_0 and conclude that the new drug is superior.

P-value gives stronger evidence in favor of H_1.

Ex5: A builder claims that heat pumps are installed in 70% of all homes being constructed today in the city of Richmond, Virginia. Would you agree with this claim if a random survey of new homes in this city showed that 8 out of 15 had heat pumps installed? Use a 0.10 level of significance.

Soln: Null hypothesis, $H_0 : p = 0.7$

Alternate hypothesis, $H_1 : p \neq 0.7$

Given : Sample size, $n = 15$

No. of homes with heat pumps, $x = 8$

Test : two-tailed test

Test statistic : Binomial variable X with $p = 0.7$ and $n = 15$

P-value : $P = 2P(X \leq 8 \text{ when } p = 0.7)$ (Since n is small)

$$= 2 \sum_{x=0}^{8} b(x; n, p)$$

$$= 2 \sum_{x=0}^{8} b(x; 15, 0.7)$$

$$= 0.2622 > 0.1$$

Decision : We donot reject H_0. Conclude that there is insufficient reason to doubt the builder's claim.

Ex 6: An automated 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.

Soln: Parameter of interest : The population variance σ^2

Null hypothesis : $H_0 : \sigma^2 = 0.01$

Alternate hypothesis : $H_1 : \sigma^2 > 0.01$

Given : $n = 20, \ s^2 = 0.0153$

Test : Upper-tailed test (χ^2-test)

Test statistic : $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$

Computation : $\chi_0^2 = \dfrac{(20-1)(0.0153)}{(0.01)^2}$

$= 29.07$

Level of significance : $\alpha = 0.05$

Critical region : $\chi_0^2 > \chi^2_{0.05, n-1}$

$\Rightarrow \chi_0^2 > \chi^2_{0.05, 19}$

$\Rightarrow \chi_0^2 > 30.14$ (from table)

Decision : Since $\chi_0^2 = 29.07 < \chi^2_{0.05,19} = 30.14$ we conclude that there is no strong evidence to reject $H_0 : \sigma^2 = 0.01$. So there is no strong evidence of a problem with incorrectly filled bottles.

Ex 7 : A manufacturer of car batteries claims that the life of the company's batteries is approximately normally distributed with a standard deviation equal to 0.9 year.

If a random sample of 10 of these batteries has a standard deviation of 1.2 years, do you think that σ > 0.9 year? Use a 0.05 level of significance.

Soln: Parameter of interest : The population variance σ^2

Null hypothesis : $H_0 : \sigma^2 = 0.9^2$

Alternate hypothesis : $H_1 : \sigma^2 > 0.9^2$

Given : $n = 10, \; s^2 = 1.2^2$

level of significance : $\alpha = 0.05$

Test : Upper-tailed test (χ^2-test)

Test statistic : $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$

Computation : $\chi_0^2 = \dfrac{(10-1)(1.2)^2}{(0.9)^2}$

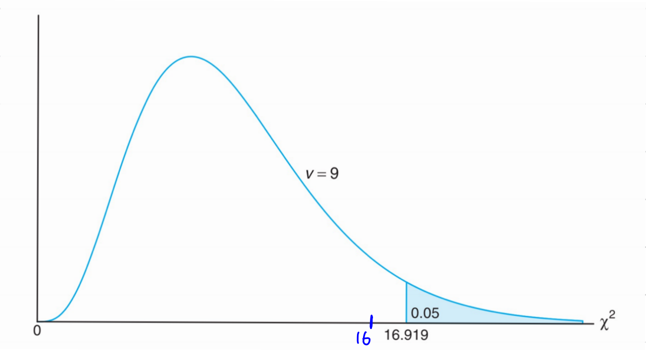
$= 16$

Critical region : $\chi_0^2 > \chi^2_{0.05,\, n-1}$

$\Rightarrow \chi_0^2 > \chi^2_{0.05,\, 9}$

$\Rightarrow \chi_0^2 > 16.919$

Decision : Since $\chi_0^2 = 16 < 16.919$, we have no sufficient evidence that $\sigma > 0.9$.



Testing for goodness-of-fit

Till now hypothesis-testing are designed for problems in which the population or probability distribution is known and the hypothesis involve the parameters of the distribution.

We now encounter another kind of hypothesis:

We donot know the underlying distribution of the population, and we wish to test the hypothesis that a particular distribution will be satisfactory as a population model.

Ex: We wish to test the hypothesis that the population is normal.

We use goodness-of-fit procedure based on the chi-square distribution.

Steps involved in the procedure:

i) Consider a random sample of size n.

ii) These n observations are arranged in a frequency histogram, having k bins or class intervals.

iii) Let O_i be the observed frequency in the ith class-interval.

iv) From the hypothesized probability distribution, we compute the expected frequency E_i, in the i^{th} class interval.

Note that each of the expected frequency is atleast equal to 5.

The test statistic is

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

The distribution of the r.v. χ^2 is approximately Chi-squared distribution with $\nu = k-1$ degree of freedom.

We reject the hypothesis if the value of χ^2 lies in the critical region, i.e, when $\chi_0^2 > \chi^2_{\alpha, k-1}$ for α level of significance.

Ex 1 : The number of defects in printed circuit boards is hypothesized to follow a Poisson distribution. A random sample of n=60 printed boards has been collected, and the following number of defects observed. Use $\alpha = 0.05$

Number of defects	Observed frequency
0	32
1	15
2	9
3	4

Soln : Given $n = 60$

We estimate the mean of the assumed Poisson distribution from the sample data

$$\lambda = \frac{0(32) + 1(15) + 2(9) + 3(4)}{60} = 0.75$$

We compute theoritical, hypothesized probability p_i as follows

$$p_1 = P(X=0) = \frac{e^{-0.75}(0.75)^0}{0!} = 0.472$$

$$p_2 = P(X=1) = \frac{e^{-0.75}(0.75)^1}{1!} = 0.354$$

$$p_3 = P(X=2) = \frac{e^{-0.75}(0.75)^2}{2!} = 0.133$$

$$p_4 = P(X \geq 3) = 1 - (p_1 + p_2 + p_3) = 0.041$$

The expected frequency $E_i = np_i$

Number of defects	Probability	Expected frequency
0	0.472	28.32
1	0.354	21.24
2	0.133	7.98
3 (or more)	0.041	2.46

Since the expected frequency in the last cell is < 3, we combine the last two cells:

Number of defects	Observed freq.	Expected freq
0	32	28.32
1	15	21.24
2 (or more)	13	10.44

Parameter of interest : The distribution of defects in printed circuit board.

Null hypothesis, H_0 : The form of distribution of defects is Poisson

Alternate hypothesis, H_1 : The form of distribution of defects is not Poisson

Test statistic : $\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$

Level of significance : $\alpha = 0.05$, reject H_0 if P-value is less than 0.05

Computation : $\chi_0^2 = \frac{(32-28.32)^2}{28.32} + \frac{(15-21.24)^2}{21.24}$

$$+ \frac{(13-10.44)^2}{10.44} = 2.94$$

Conclusion : From table $\chi^2_{0.05,2} = 5.99$ and $\chi_0^2 = 2.94 < 5.99$. Thus P-value is greater that 0.05. We fail to reject H_0. The distribution of defects is Poisson.