STATISTICS

Module 2 *of*
MATHEMATICS AND STATISTICS

Marina Santacroce

marina.santacroce@unibocconi.it
Department of Decision Sciences, Bocconi University

L17: Hypothesis testing.

References: S.Ross *Introduction to* Probability and Statistics for Engineers and Scientists: 8.1, 8.2, 8.3, 8.4, 8.5, 8.6

Outline

Introduction

Hypothesis testing in a normal population
- Hypothesis testing for the mean with known and unknown variance
- Hypothesis testing for the variance
- Hypothesis testing for the difference in means of two normal populations with known and unknown variance

Hypothesis testing in a Bernoulli population
- Hypothesis testing for the mean in a Bernoulli population

Hypothesis testing

The goal of hypothesis testing is to use the data of a random sample to test some particular hypotheses on the population parameters.

- A typical example is a firm that will invest in a new machine if the probability that an item produced by the new machine is defective is less than 0.01.
- A statistical hypothesis is a statement about the parameters of the population of interest.
- Hypothesis testing is a *procedure* for determining whether or not the values of a random sample are consistent with the hypothesis.

- $X_1, \ldots X_n$, are sampled from a population with distribution F_θ
- F_θ is specified up to a set of *unknown parameters* θ

$\Rightarrow$ use the sample to test a specific hypothesis about θ.

The hypothesis to be tested H_0 is called *null hypothesis*. It can be:

- *simple* (hypothesis) ($H_0 : \theta = \theta_0$) H_0 completely specifies the distribution;
- *composite* (hypothesis) ($H_0 : \theta \leq \theta_0$).

Hypothesis testing

Given $X_1, \ldots X_n$, a testing procedure for H_0 decides if to accept H_0 or not:

$\Rightarrow$ the *critical region C* of the space where $X_1, \ldots X_n$ take values:

- if the random sample $X_1, \ldots X_n \notin C \Rightarrow H_0$ is accepted;
- if the random sample $X_1, \ldots X_n \in C \Rightarrow H_0$ is rejected.

The testing procedure is subject to two types of errors:

- the test *rejects* H_0, when it is indeed *true* (type I error)
- the test *accepts* H_0, when it is *false* (type II error).

$\Rightarrow$ The test is built is such a way that it will reject H_0 if the data are *very unlikely* when H_0 is true.

$\Rightarrow$ We fix the key quantity α, suitably small, and require that

if H_0 is true the probability of rejecting it is less or equal to α.

If $H_0 : \theta \in w_0$, the **main idea** is **to find** the **critical region** C **using the distribution of an estimator of** θ $d(X_1, \ldots X_n)$:

- $C = \{\text{reject } H_0\} \Leftrightarrow d(X_1, \ldots X_n)$ is *far away from* w_0;
- under H_0 the probability of rejecting it has to be less or equal to α.

Hypothesis testing for the mean of a normal population (with known variance)

$X_1, \dots X_n$ is a random sample from a normal population with unknown mean μ and known variance σ^2. Given $X_1, \dots X_n$,

- the goal is to *test the null (simple) hypothesis*

$$H_0 : \mu = \mu_0$$

  against the alternative hypothesis

$$H_1 : \mu \neq \mu_0$$

  where μ_0 is some specified constant.

- Chosen a significance level α, with $\alpha \in (0, 1)$, we determine the critical region C of the test by using the distribution of $\overline{X}$:

$\Rightarrow$ if $\mu = \mu_0$, we find c such that the probability of rejecting H_0 is equal to α

$$\mathbb{P}_{\mu_0}(|\overline{X} - \mu_0| > c) = \alpha.$$

If $\mu = \mu_0$, $\quad \overline{X} \sim \mathrm{N}(\mu_0, \frac{\sigma^2}{n})$, $\Rightarrow \mathbb{P}_{\mu_0}(|\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}| > z_{\frac{\alpha}{2}}) = \alpha.$

$\Rightarrow$ Therefore, $\quad c = z_{\frac{\alpha}{2}} \sigma / \sqrt{n}$ and the critical region is

$$C = \{X_1, \dots X_n : |\overline{X} - \mu_0| > z_{\frac{\alpha}{2}} \sigma / \sqrt{n}\}$$

Example: hypothesis testing on μ (known variance)

Example: Consider a signal whose value is an unknown constant μ which is transmitted from a location A to a location B We assume the signal received at B is $X = \mu + Y$ where the noise Y is normal with mean 0 and variance 4. Suppose today it has been transmitted 5 times and that $\overline{x} = 9.5$.
We want to test the null hypothesis

$$H_0 : \mu = 8 \qquad \text{against the alternative} \qquad H_1 : \mu \neq 8$$

Since $\mu_0 = 8$ and $z_{0.025} = 1.96$, the critical region for $\alpha = 0.05$

$$C = (-\infty, 8 - 1.96\frac{2}{\sqrt{5}}) \cup (8 + 1.96\frac{2}{\sqrt{5}}, \infty) = (-\infty, 6.25) \cup (9.75, \infty).$$

Therefore, *given the observed value is* $\overline{x} = 9.5$*, the null hypothesis* $\mu = 8$ *is accepted with significance level* $\alpha = 0.05$.

Note that for $\alpha = 0.1$, since $z_{0.05} = 1.645$ and

$$C = (-\infty, 8 - 1.645\frac{2}{\sqrt{5}}) \cup (8 + 1.645\frac{2}{\sqrt{5}}, \infty) = (-\infty, 6.53) \cup (9.47, \infty).$$

Thus, $\overline{x} = 9.5 \in C$ and the null hypothesis would be rejected.

p-value of the test

Example: Consider a signal whose value is an unknown constant μ which is transmitted from a location A to a location B. We assume the signal received at B is $X = \mu + Y$ where the noise Y is normal with mean 0 and variance 4.
Suppose today the signal has been transmitted 5 times and that $\overline{x} = 9.5$.
We want to test the null hypothesis

$$H_0 : \mu = 8 \qquad \text{versus} \qquad H_1 : \mu \neq 8$$

Note that given $\overline{x} = 9.5$ has been observed,

$$\text{from} \quad |9.5 - 8| = z_{\frac{\alpha}{2}} 2/\sqrt{5} \quad \Rightarrow \quad \alpha = 0.094$$

we find the *p*-value of the test p=0.094, that is the *critical significance level* in the sense that we find the minimum value of α such that, given the observed statistics, the test will be rejected:

- for any significance level $\alpha < p$ the null hypothesis will be accepted;
- for any significance level $\alpha \geq p$ the null hypothesis will be rejected.

Probability of type II error and power-function of the test

Recall the alternative hypothesis

$$H_1 : \mu \neq \mu_0.$$

The *probability* of the type II error[1] depends on the value μ, with $\mu \neq \mu_0$, and is

$$\beta(\mu) = \mathbb{P}_\mu(\text{accept } H_0) = \mathbb{P}_\mu(|\overline{X} - \mu_0| \leq z_{\frac{\alpha}{2}}\sigma/\sqrt{n}), \quad \mu \neq \mu_0.$$

The function $1 - \beta(\mu)$ is called power-function of the test which is the probability of rejection when μ, with $\mu \neq \mu_0$, is the true value.

A desirable test has:

- the level of significance α small
- the power function $1 - \beta(\mu)$ big

[1] accepting the null hypothesis when it is false

One-sided tests for the mean of a normal population (with known variance)

$X_1, \ldots X_n$ is a random sample from a normal population with unknown mean μ and known variance σ^2. Given $X_1, \ldots X_n$,

- the goal is to *test the null (simple) hypothesis*

$$H_0 : \mu = \mu_0$$

against the alternative hypothesis

$$H_1 : \mu > \mu_0$$

where μ_0 is some specified constant.

- Chosen a significance level α, we determine the critical region C of the test taking into account the distribution of $\overline{X}$ and that small values of $\overline{X}$ are more likely under H_0 (rather than under H_1):

$\Rightarrow$ if $\mu = \mu_0$, we find c such that the probability of rejecting H_0 is equal to α

$$\mathbb{P}_{\mu_0}(\overline{X} - \mu_0 > c) = \alpha.$$

If $\mu = \mu_0$, $\quad \overline{X} \sim \mathrm{N}(\mu_0, \frac{\sigma^2}{n})$, $\Rightarrow \mathbb{P}_{\mu_0}(\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha) = \alpha.$

$\Rightarrow$ Therefore, $\quad c = z_\alpha \sigma/\sqrt{n} \quad$ and $\quad C = \{X_1, \ldots X_n : \overline{X} > \mu_0 + z_\alpha \sigma/\sqrt{n}\}$

One-sided tests for the mean of a normal population (with known variance)

Under the same hypotheses, similarly, one can consider the test

$$H_0 : \mu = \mu_0$$

against the alternative hypothesis

$$H_1 : \mu < \mu_0,$$

and compute the critical region $C = \{X_1, \dots X_n : \overline{X} < \mu_0 - z_\alpha \sigma/\sqrt{n}\}$

- the test rejects H_0 if $\overline{X} < \mu_0 - z_\alpha \sigma/\sqrt{n}$
- the test accepts H_0 if $\overline{X} \geq \mu_0 - z_\alpha \sigma/\sqrt{n}$.

Moreover, the power-function[2] of the test $1 - \beta(\mu)$, is

$$1 - \beta(\mu) = \mathbb{P}_\mu(\overline{X} < \mu_0 - z_\alpha \sigma/\sqrt{n}) = \Phi(-z_\alpha + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma}), \quad \mu < \mu_0.$$

$1 - \beta(\mu)$ decreases in μ and $1 - \beta(\mu_0) = \Phi(-z_\alpha) = 1 - \Phi(z_\alpha) = \alpha$

$\alpha \leq 1 - \beta(\mu)$: the probability to reject H_0 is higher if H_0 is false rather than if it is true;

$\beta(\mu) \leq 1 - \alpha$: the probability to accept H_0 is higher if H_0 is true rather than if it is false.

- Note that the tests $H_0 : \mu = \mu_0$ or $H_0 : \mu \geq \mu_0$ against $H_1 : \mu < \mu_0$ are equivalent.

[2] probability of rejection when μ is the true value

Hypothesis testing for the mean of a normal population (with unknown variance)

Let $X_1, \dots X_n$ be a random sample from a normal population with *unknown* mean μ and *unknown* variance σ^2.
We want to test *the null (simple) hypothesis*

$$H_0 : \mu = \mu_0$$

against the alternative hypothesis

$$H_1 : \mu \neq \mu_0.$$

- Chosen a significance level α, we determine the critical region C using that, under H_0, $\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$ is a t with $n-1$ degrees of freedom:
  we find c such that the probability of the type I error is α:

$$\mathbb{P}_{\mu_0}(|\overline{X} - \mu_0| > c) = \alpha \quad \text{and} \quad \mathbb{P}_{\mu_0}\left(\left|\frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}}\right| > t_{\frac{\alpha}{2},n-1}\right) = \alpha.$$

$$\Rightarrow \qquad c = t_{\frac{\alpha}{2},n-1} S/\sqrt{n} \quad \text{and} \quad C = \{X_1, \dots X_n : |\overline{X} - \mu_0| > t_{\frac{\alpha}{2},n-1} S/\sqrt{n}\}$$

- the test *accepts* H_0 if $\overline{X} \in [\mu_0 - t_{\frac{\alpha}{2},n-1} S/\sqrt{n}, \mu_0 + t_{\frac{\alpha}{2},n-1} S/\sqrt{n}]$.
- the test *rejects* H_0 otherwise.

One-sided tests for the mean of a normal population (with unknown variance)

Let $X_1, \dots X_n$ be a random sample from a normal population with *unknown* mean μ and *unknown* variance σ^2.
We want to *test the null (simple) hypothesis*

$$H_0 : \mu = \mu_0$$

against the alternative hypothesis

$$H_1 : \mu > \mu_0.$$

- Chosen a significance level α, we determine the critical region C using that, under H_0, $\frac{\overline{X} - \mu_0}{S/\sqrt{n}}$ is a t with $n-1$ degrees of freedom:
  we find c such that the probability of the type I error is α:

$$\mathbb{P}_{\mu_0}(\overline{X} - \mu_0 > c) = \alpha \quad \text{and} \quad \mathbb{P}_{\mu_0}\left(\frac{\overline{X} - \mu_0}{\frac{S}{\sqrt{n}}} > t_{\alpha,n-1}\right) = \alpha.$$

$$\Rightarrow \qquad c = t_{\alpha,n-1} S/\sqrt{n} \quad \text{and} \quad C = \{X_1, \dots X_n : \overline{X} > \mu_0 + t_{\alpha,n-1} S/\sqrt{n}\}$$

- the test *accepts* H_0 if $\overline{X} \leq \mu_0 + t_{\alpha,n-1} S/\sqrt{n}$,
- the test *rejects* H_0 otherwise.

- Similarly, testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$,
  - the test *accepts* H_0 if $\overline{X} \geq \mu_0 - t_{\alpha,n-1} S/\sqrt{n}$,
  - the test *rejects* H_0 otherwise.

Example: hypothesis testing on μ (unknown variance)

Example: Among patients having medium-high blood cholesterol levels 50 volunteers were recruited to test a new drug to reduce cholesterol. After one month of new drug treatment the changes in their blood cholesterol were noted. If the average change is $\overline{x} = -14.8$ with a sample standard deviation of $s = 6.4$, what conclusions can be drawn?

- $H_0 : \mu = 0$ (no effect) against $H_1 : \mu < 0$ (reduction due to the drug)
- $n = 50$ thus the t distribution of $\overline{X}\sqrt{n}/S$ is a normal (under H_0);
- $C = \{\overline{X} < -z_\alpha S/\sqrt{n}\} = \{\overline{X} < -z_\alpha 6.4/\sqrt{50}\}$, for $\alpha = 0.1$ $C = \{\overline{X} < -1.163\}$, $\alpha = 0.0001$ $C = \{\overline{X} < -3.276\}$ ...
- in fact the p value for the test is zero:

$$z_\alpha = 14.8\frac{\sqrt{50}}{6.4} = 16.35 \Rightarrow \alpha = 0$$

The data are not consistent with H_0.

Hypothesis testing for the mean of a population with known and unknown variance

- If the sample is reasonably large, *exploiting the CLT*
- the illustrated tests procedures can be used for testing hypothesis on the mean of populations which are *not normal*.

Hypothesis testing for the variance of a normal population

Let $X_1, \dots X_n$ be a random sample from a normal population with *unknown* mean μ and *unknown* variance σ^2. We want to test *the null hypothesis*

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

against the alternative hypothesis

$$H_1 : \sigma^2 \neq \sigma_0^2.$$

- Given the significance level α, we determine the critical region C using that, under H_0, $(n-1)\frac{S^2}{\sigma_0^2}$ is a χ^2 with $n-1$ degrees of freedom, since

$$\mathbb{P}_{\sigma_0^2}\left(\chi^2_{1-\frac{\alpha}{2},n-1} \leq (n-1)\frac{S^2}{\sigma_0^2} \leq \chi^2_{\frac{\alpha}{2},n-1}\right) = 1-\alpha$$

with $\chi^2_{\frac{\alpha}{2},n-1}$ such that $\mathbb{P}(\chi^2_{n-1} > \chi^2_{\frac{\alpha}{2},n-1}) = \frac{\alpha}{2}$.

Therefore,

- the test *accepts* H_0 if $S^2 \in \left[\frac{\sigma_0^2\chi^2_{1-\frac{\alpha}{2},n-1}}{n-1}, \frac{\sigma_0^2\chi^2_{\frac{\alpha}{2},n-1}}{n-1}\right]$
- the test *rejects* H_0 otherwise.

Hypothesis testing for the equality of the variances of two populations

Let $X_1, \ldots X_n$ and $Y_1, \ldots Y_m$ be two random samples from two normal populations with *unknown* means and *unknown* variances σ_1^2 and σ_2^2. We want to test the equality of the variances in the two populations:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

versus the alternative

$$H_1 : \sigma_1^2 \neq \sigma_2^2.$$

- Given α, we use that $(n-1)\frac{S_1^2}{\sigma_1^2}$ and $(m-1)\frac{S_2^2}{\sigma_2^2}$ are two independent χ^2, respectively with $n-1$ and $m-1$ degrees of freedom. Under H_0, $\frac{S_1^2}{S_2^2}$ has the F-distribution with $n-1$ and $m-1$ degrees of freedom. Therefore,

$$\mathbb{P}\left(F_{1-\frac{\alpha}{2},n-1,m-1} \leq \frac{S_1^2}{S_2^2} \leq F_{\frac{\alpha}{2},n-1,m-1}\right) = 1-\alpha$$

where $F_{\frac{\alpha}{2},n-1,m-1}$ is such that $\mathbb{P}(F_{n-1,m-1} > F_{\frac{\alpha}{2},n-1,m-1}) = \frac{\alpha}{2}$.

Therefore,

- the test *accepts* H_0 if $\frac{S_1^2}{S_2^2} \in [F_{1-\frac{\alpha}{2},n-1,m-1}, F_{\frac{\alpha}{2},n-1,m-1}]$
- the test *rejects* H_0 otherwise.

Hypothesis testing for the difference in means of two normal populations

Let $X_1, \ldots X_n$ and $Y_1, \ldots Y_m$ be two random samples from two normal populations with *unknown* means μ_1 and μ_2 and *known* variances σ_1^2 and σ_2^2. We want to test the equality of the means in the two populations:

$$H_0 : \mu_1 = \mu_2$$

versus the alternative

$$H_1 : \mu_1 \neq \mu_2.$$

- Fixed α we determine the critical region using that $\overline{X} - \overline{Y} \sim \mathrm{N}(\mu_1 - \mu_2, \sigma_1^2/n + \sigma_2^2/m)$. Under H_0,

$$\mathbb{P}\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X} - \overline{Y}}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}} \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

Therefore,

- the test *accepts* H_0 if

$$\overline{X} - \overline{Y} \in \left[-z_{\frac{\alpha}{2}}\sqrt{\sigma_1^2/n + \sigma_2^2/m}, z_{\frac{\alpha}{2}}\sqrt{\sigma_1^2/n + \sigma_2^2/m}\right]$$

- the test *rejects* H_0 otherwise.

Hypothesis testing for the difference in means of two normal populations

Let $X_1, \dots X_n$ and $Y_1, \dots Y_m$ be two random samples from two normal populations with *unknown* means μ_1 and μ_2 and *unknown* (common) variance σ^2. We want to test the equality of the means in the two populations:

$$H_0 : \mu_1 = \mu_2$$

versus the alternative

$$H_1 : \mu_1 \neq \mu_2.$$

$S_p^2 = \frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}$ denotes the pooled variance. Fixed α we determine the critical region using that

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}.$$

Under H_0,

$$\mathbb{P}\Big(- t_{\frac{\alpha}{2},n+m-2} \leq \frac{\overline{X} - \overline{Y}}{S_p\sqrt{1/n + 1/m}} \leq t_{\frac{\alpha}{2},n+m-2}\Big) = 1 - \alpha.$$

Therefore,

- the test *accepts* H_0 if

$$\overline{X} - \overline{Y} \in [-t_{\frac{\alpha}{2},n+m-2} S_p\sqrt{1/n + 1/m}, t_{\frac{\alpha}{2} n+m-2} S_p\sqrt{1/n + 1/m}]$$

- the test *rejects* H_0 otherwise.

Hypothesis testing for the mean in a Bernoulli population

Suppose to observe a random sample $X_1, \ldots, X_n$ from a Bernoulli population with unknown parameter p and to denote $X = \sum_{i:1}^{n} X_i$. We want to test

$$H_0 : p \leq p_0$$

versus the alternative

$$H_1 : p > p_0.$$

- X is a binomial with parameters n and p and the the probability X assumes large values is increasing in p;
- the critical region $C = \{X \geq k^*\}$ where k^* is defined as the smallest value k such that $\mathbb{P}_{p_0}(X \geq k) \leq \alpha$. Once k^* has been found
  - the test *rejects* H_0 if $X \geq k^*$
  - the test *accepts* H_0 otherwise.
- if n is large, by the CLT, under H_0, X is approximately normal and $\frac{X-np_0}{\sqrt{np_0(1-p_0)}} \approx\sim \mathrm{N}(0,1)$

$$\mathbb{P}\left(\frac{X - np_0}{\sqrt{np_0(1-p_0)}} \leq z_\alpha\right) = 1 - \alpha$$

  from which we obtain
  - the test *rejects* H_0 if $X \geq np_0 + z_\alpha \sqrt{np_0(1-p_0)}$
  - the test *accepts* H_0 otherwise.

Example: hypothesis testing for the mean in a Bernoulli population

Example: A computer chip manufacturer claims that no more than 2 percent of the chips it sends are defective. An electronics company, impressed with this claim, has purchased a large quantity of such chips. The company decides to test a sample of 300 chips to determine if the claim can be trusted. If 10 of these chips are defective, should the claim be rejected?

We want to test with $\alpha = 0.05$,

$$H_0 : p \leq 0.02$$

versus

$$H_1 : p > 0.02.$$

- the critical region $C = \{X \geq k^*\}$ where k^* is defined as the smallest k such that

  $$X \sim \text{bin}(300, 0.02) \qquad \mathbb{P}(X \geq k) \leq 0.05.$$

  Since, $\mathbb{P}(X \geq 10) = 0.0818$, $\mathbb{P}(X \geq 9) = 0.0409 \Rightarrow k^* = 9$ and $C = \{X \geq 9\}$.
- Thus, given we found 10 defective chips, 10 falls in the critical region and the test rejects H_0 with $\alpha = 0.05$.
- Since $n = 300$ is large, we can approximate X with a normal and $\frac{X-6}{\sqrt{6(0.98)}} \approx\sim \mathrm{N}(0,1)$ and $z_{0.05} = 1.645$

$\Rightarrow$ the test *rejects* H_0 since $10 \geq 6 + 1.645\sqrt{6(0.98)} = 9.9889.$