Foundations of probability theory

Joaquín Rapela

Gatsby Computational Neuroscience Unit
University College London

December 28, 2023

Contents

Contents

Course logistics

Contents

Working examples

Example 1

The average running speed of control mice is 1 cm/sec (cohort size 100 mice). The sample average running speed of an equally-sized cohort of transgenic mice is $\bar{x} = 2.7\,\text{cm/sec}$ and the sample standard deviation is $s = 10$ cm/sec. Is the average running speed of the transgenic cohort larger than that of control mice?

Working examples

Example 2

We want to study the effect of a new drug on visual electrophysiology in humans. In a control condition we observe that the mean peak (averaged across a cohort of 50 partcipants) evoked response potential (ERP) over V1 during the first 200 ms after stimuli presentation is 2 mV. The sample mean peak ERP for medicated subjects is $\bar{x} = 1.3\,\text{mV}$ and the sample standar deviation is $s = 2.6$ mV. We want to know if taking the new drug changes the mean evoked ERP over V1.

Statistical remarks

1. It is frequent in statistics to assume that observed data follows a given probability distribution. For example a:
   - normal distribution with parameters mean μ and variance σ^2, $\mathcal{N}(\mu, \sigma^2)$,
   - Poisson distribution with expected rate parameter λ, $\mathcal{P}(\lambda)$,
   - Binomial distribution with number of observation parameter n and with a success probability parameter p, $\mathcal{B}(n, p)$.

Statistical remarks

2. One branch of statistics, **estimation theory**, provides tools to estimate parameters of distributions from observations.
3. Another branch of statistics, **hypothesis testing**, provides tools to make statistically-informed decisions about values of parameters of distributions.

Statistical remarks

4. To estimate parameters, or to make decisions about them, we use observations, $x_1, \ldots, x_N$, that are **independent and identically distributed**.

Statistical remarks

5. To estimate parameters, or to make decisions about them, we use observations, $x_1, \ldots, x_N$, that are **independent and identically distributed**.

Example 1

An observation will be the average speed of a transgentic mouse during an experimetal session. We assume that the average speeds of all mice are samples from a common probability density function (identically distributed) and that average speeds are independent across mice (independent).

Statistical remarks

6. To estimate parameters, or to make decisions about them, we use observations, $x_1, \ldots, x_N$, that are **independent and identically distributed**.

Example 1

An observation will be the average speed of a transgentic mouse during an experimetal session. We assume that the average speeds of all mice are samples from a common probability density function (identically distributed) and that average speeds are independent across mice (independent).

Example 2

An observation will be the peak ERP of a medicated cohort subject. We assume tha these ERPs are samples from the same probability density function (identically distributed) and that these ERPs are independent across subjects (independent).

Statistical remarks

7. A goal of statistics is to **infer properties of the population** (e.g., the effect of the genetic manipulation on the running speed of mice) from **properties of the sample** (e.g., the effect of the manipulation on the running speed of the 100 sampled mice).

Statistical remarks

Theorem (Central Limit Theorem)

Let $X_1, \ldots, X_N$ be independent and identically distributed random variables with mean μ and variance σ^2. Let N be sufficiently large. Then $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{N})$.

Note: if σ^2 is unknown, for large N, we can estimate σ^2 with the sample variance $s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$. Then $\bar{X}$ is approximately distributed as $\mathcal{N}(\mu, \frac{s^2}{N})$.

Contents

Null and alternative hypothesis

- In hypothesis testing we work with a **null hypothesis**, $\mathcal{H}_0$, and an **alternative hypothesis**, $\mathcal{H}_a$, collect a sample of data $x_1, \ldots, x_N$, and test if this data provides sufficient statistical evidence in favor of the alternative hypothesis. If this happens we reject the null hypothesis.
- However, if the collected data does not provide sufficient statistical evicence in favor of the aternative hypothesis, we do not accept the null hypothesis, but we say that we failed to reject it. **Hypothesis tests do not prove null hypothesis, they only provide statistical evidence to reject it, or fail to reject it.**

Null and alternative hypothesis

- In hypothesis testing we work with a **null hypothesis**, $\mathcal{H}_0$, and an **alternative hypothesis**, $\mathcal{H}_a$, collect a sample of data $x_1, \dots, x_N$, and test if this data provides sufficient statistical evidence in favor of the alternative hypothesis. If this happens we reject the null hypothesis.
- However, if the collected data does not provide sufficient statistical evicence in favor of the aternative hypothesis, we do not accept the null hypothesis, but we say that we failed to reject it. **Hypothesis tests do not prove null hypothesis, they only provide statistical evidence to reject it, or fail to reject it.**

The hypothesis that we aim to prove should be the alternative one.

Null and alternative hypothesis

Example 1

Is the average running speed of the transgenic cohort larger than that of the control cohort (i.e., 2 cm/sec)?

$\mathcal{H}_0$: the average running speed of the trangenic cohort is 2 cm/sec.

$\mathcal{H}_a$: the average running speed of the transgenic cohort is lager than 2 cm/sec.

Null and alternative hypothesis

Example 2

Is the mean peak visual ERP in the first 200 ms post stimuli different in medicated than in control subjects (i.e., 2 mV)?

$\mathcal{H}_0$: the mean peak visual ERP in medicated subjects is 2 mV.

$\mathcal{H}_a$: the mean peak visual ERP in medicated subjects is different from 2 mV.

One- and two-tailed tests of hypothesis

One-tailed test of hypothesis directionality is suggested by the alternative hypothesis.

Example 1

Is is a one-tailed hypothesis test because the alternative hypothesis requires that the mean speed of the transgenic mice be larger (directionality) than that of the control mice.

One- and two-tailed tests of hypothesis

One-tailed test of hypothesis directionality is suggested by the alternative hypothesis.

Example 1

Is is a one-tailed hypothesis test because the alternative hypothesis requires that the mean speed of the transgenic mice be larger (directionality) than that of the control mice.

Two-tailed test of hypothesis directionality is not suggested by the alterntive hypothesis.

Example 2

It is a two-tailed hypothesis test because the alternative hypothesis requires that the visual ERP of the medicated subjects be different (no directionality) than that of the control subjects.

Test statistic and its sampling distribution

- To perform a hypothesis test we propose a **test statistic**, a function of the sample data, like the sample mean:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{1}$$

- Because the sample data is random, the test statistic is also random. To perform hypothesis tests we need to know the distribution of the test statistic, which is called the **sampling distribution**.
- If the observed test statistic is highly unprobable under $\mathcal{H}_0$ (based on the sampling distribution), and consistent with $\mathcal{H}_a$, we reject the null hypothesis.

Test statistic and its sampling distribution

Because both examples are tests for the population mean, μ, because the sample mean, $\bar{x}$, is a good estimator of the population mean, and because both examples use a large number of samples, we will use the standardized sample mean as our test statistic:

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{N}}$$

From the central limit theorem, we know the sampling distribution of this test statistic:

$$Z \sim \mathcal{N}(0, 1)$$

Rejection and non-rejection regions

The **reject region** is a region of low probability under the null hypohesis, which is consitent with alternative hypothesis. The **non-reject region**, is a region of large probability under the null hypothesis, that is inconsistent with the alternative hypothesis.

Rejection and non-rejection regions

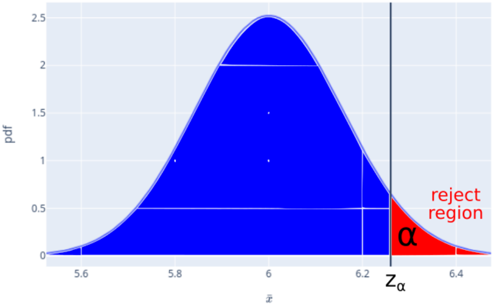
Example 1

We want to reject the null hypothesis if the sample mean velocity is larger than 2 cm/sec (the population mean under the null hypothesis). That is, we want to reject the null hypothesis if z is large and positive. How large is large?

To answer this question we define the **Type I error** of a test, as the error of rejecting the null hypothesis when it is valid. We also define the **significance level of the hypothesis test**, α, as the probability of Type I error admitted by the test.

When designing a test we first decide on its significane level α. We then reject the null hypothesis if z is larger than the value z_α that leaves α probability to its right, We call this value the **critical value** of the test (see figure on next slide).

Two types of errors, confidence leval and p-value



Type I error : reject the null hypothesis when it is true.
Type II error : not reject the null hypothesis when the alternative one is true.

Some statistical tests are designed to constrain the probability

Steps to perform a hypothesis test

We start assuming that we know the distribution of the experimental samples and we agree on a sample size N (in Section ?? we discuss how to select and optimal sample size). We set the confidence level α of the test (i.e., the probability that the test statistic falls in the reject region given that the null hypothesis is valid).

1. collect an experimental sample $x_1, \ldots, x_N$.
2. compute the value of the test statistic corresponding to the collected experimental sample (e.g., sample mean, Eq. 1)
3. calculate the sample statistic, t_{obs}.
4. for testing based on confidence level:
   a. divide the space of all possible values of the test statistic on a reject and a non-reject regons at a confidence level α.
   b. if the value of the test statistic falls in the reject region, reject the null hypothesis at a confidence α. If it does not, do not reject the null hypothesis at this confidence.

Central limit theorem