Class 18 in-class problems, 18.05, Spring 2022

Concept questions

Concept question 1. NHST

You collect data from an experiment and do a left-sided z-test with significance 0.1. You find the z-value is 1.8

(i) Which of the following computes the critical value for the rejection region?

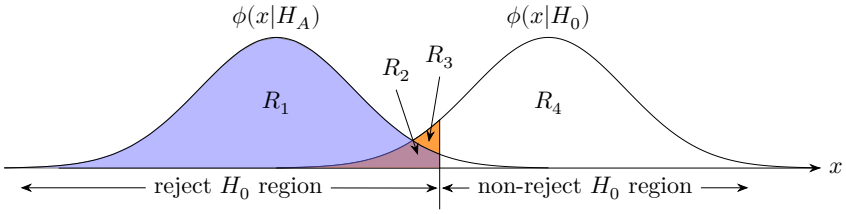
(a) `pnorm(0.1, 0, 1)` (b) `pnorm(0.9, 0, 1)`
(c) `pnorm(0.95, 0, 1)` (d) `pnorm(1.8, 0, 1)`
(e) `1 - pnorm(1.8, 0, 1)` (f) `qnorm(0.05, 0, 1)`
(g) `qnorm(0.1, 0, 1)` (h) `qnorm(0.9, 0, 1)`
(i) `qnorm(0.95, 0, 1)`

(ii) Which of the above computes the p-value for this experiment?

(iii) Should you reject the null hypothesis?

Concept question 2. Power

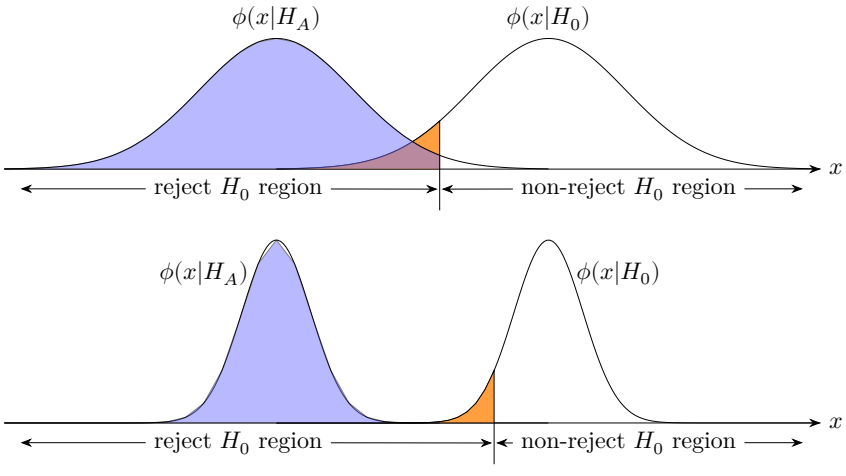
The power of the test in the graph is given by the area of



(a) R_1 (b) R_2 (c) $R_1 + R_2$ (d) $R_1 + R_2 + R_3$

Concept question 3. Higher power

Which of the tests below has higher power?



(1) Top graph (2) Bottom graph

Board questions

Problem 1. Significance level and power

Our data x follows a binomial(θ, 10) distribution with θ unknown.

The rejection region is boxed in orange. The corresponding probabilities for different hypotheses are shaded below it.

x	0	1	2	3	4	5	6	7	8	9	10
$H_0: p(x\|\theta = 0.5)$	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	.001
$H_A: p(x\|\theta = 0.6)$	0.000	0.002	0.011	0.042	0.111	0.201	0.251	0.215	0.121	0.040	0.006
$H_A: p(x\|\theta = 0.7)$	0.000	0.000	0.001	0.009	0.037	0.103	0.200	0.267	0.233	0.121	0.028

(a) Find the significance level of the test.

(b) Find the power of the test for each of the two alternative hypotheses.

(c) What is the probability of a type I error? type II?

Problem 2. z and one-sample t-test

For both problems use significance level $\alpha = 0.05$.

Assume the data 2, 4, 4, 10 are independently drawn from a $N(\mu, \sigma^2)$.

The hypotheses are: H_0: $\mu = 0$ and H_A: $\mu \neq 0$.

(a) Is the test one or two-sided? If one-sided, which side?

(b) Assume $\sigma^2 = 16$ is known and test H_0 against H_A.

(c) Now assume σ^2 is unknown and test H_0 against H_A.

Problem 3. Two-sample t-test

Real data from 1408 women admitted to a maternity hospital for (i) medical reasons or through (ii) unbooked emergency admission. The duration of pregnancy is measured in complete weeks from the beginning of the last menstrual period.

Medical: 775 obs. with $\bar{x} = 39.08$ and $s^2 = 7.77$.

Emergency: 633 obs. with $\bar{x} = 39.60$ and $s^2 = 4.95$

(a) Set up and run a two-sample t-test to investigate whether the duration differs for the two groups.

(b) What assumptions did you make?

Discussion questions

1. Significance and power

The null distribution for test statistic x is $N(4, 8^2)$. The rejection region is $\{x \geq 20\}$.

What is the significance level and power of this test?

2. Type I errors Q1

Suppose a journal will only publish results that are statistically significant at the 0.05 level. What percentage of the papers it publishes contain type I errors?

3. Type I errors Q2

Jerry desperately wants to cure diseases but he is terrible at designing effective treatments. He is however a careful scientist and statistician, so he randomly divides his patients into control and treatment groups. The control group gets a placebo and the treatment group gets the experimental treatment. His null hypothesis H_0 is that the treatment is no better than the placebo. He uses a significance level of $\alpha = 0.05$. If his p-value is less than α he publishes a paper claiming the treatment is significantly better than a placebo.

(a) Since his treatments are never, in fact, effective what percentage of his experiments result in published papers?

(b) What percentage of his published papers contain type I errors, i.e. describe treatments that are no better than placebo?

4. Type I errors Q3

Jen is a genius at designing treatments, so all of her proposed treatments are effective. She is also a careful scientist and statistician, so she too runs double-blind, placebo controlled, randomized studies. Her null hypothesis is always that the new treatment is no better than the placebo. She also uses a significance level of $\alpha = 0.05$ and publishes a paper if $p < \alpha$.

(a) How could you determine what percentage of her experiments result in publications?

(b) What percentage of her published papers contain type I errors, i.e. describe treatments that are, in fact, no better than placebo?

18.05 Introduction to Probability and Statistics
Spring 2022