Definition: Power

Setup: Θ parameter space, $\Theta_0 \subset \Theta$, $\Theta_1 = \Theta \setminus \Theta_0$. Consider

$$H_0 : \theta \in \Theta_0 \text{ v.s. } H_1 : \theta \in \Theta_1$$

Suppose we have some test for this hypothesis.
The *power function* is defined as the mapping

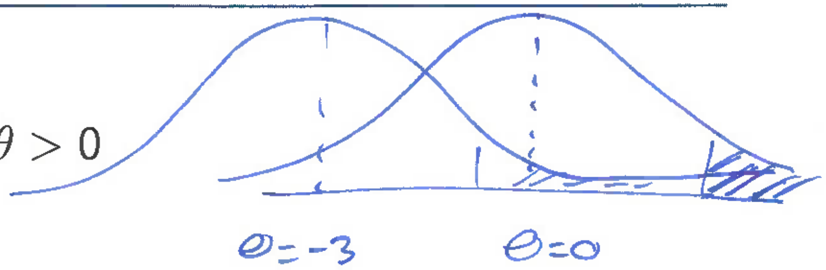
$$\beta : \Theta \to [0,1], \beta(\theta) = \underline{P_\theta(\text{reject } H_0)} = P_\theta(Y \in R)$$

$\theta \in \Theta$

If $\theta \in \Theta_0$ then we want $\beta(\theta)$ to be small.
If $\theta \in \Theta_1$ then we want $\beta(\theta)$ to be large.

Example: $X \sim N(\theta, 1)$, $\theta \in \mathbb{R}$ unknown

$$H_0 : \theta \leq 0 \quad \text{against} \quad H_1 : \theta > 0$$



Level α test

$$\Theta = \mathbb{R}, \Theta_0 = (-\infty, 0], \Theta_1 = (0, \infty)$$

Rejection region

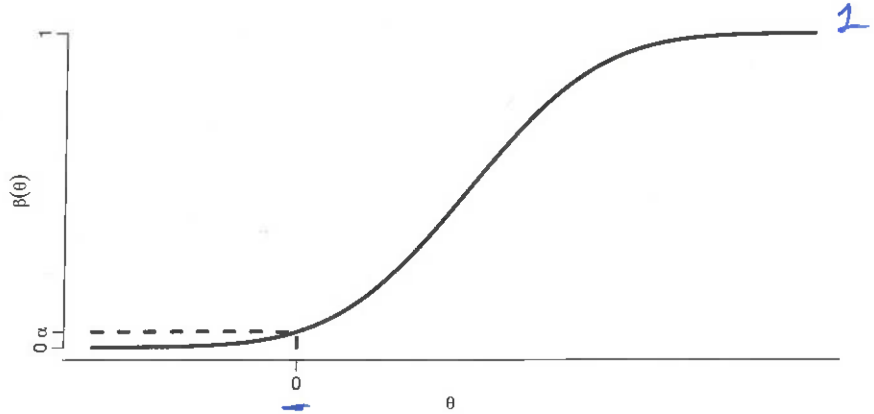
$$\boxed{R = [c, \infty)}$$

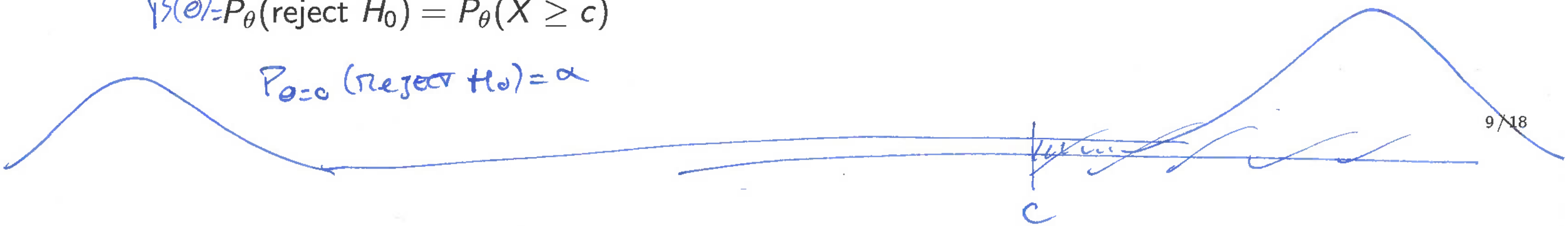
Choose c s.t. $\Phi(c) = 1 - \alpha$. Then

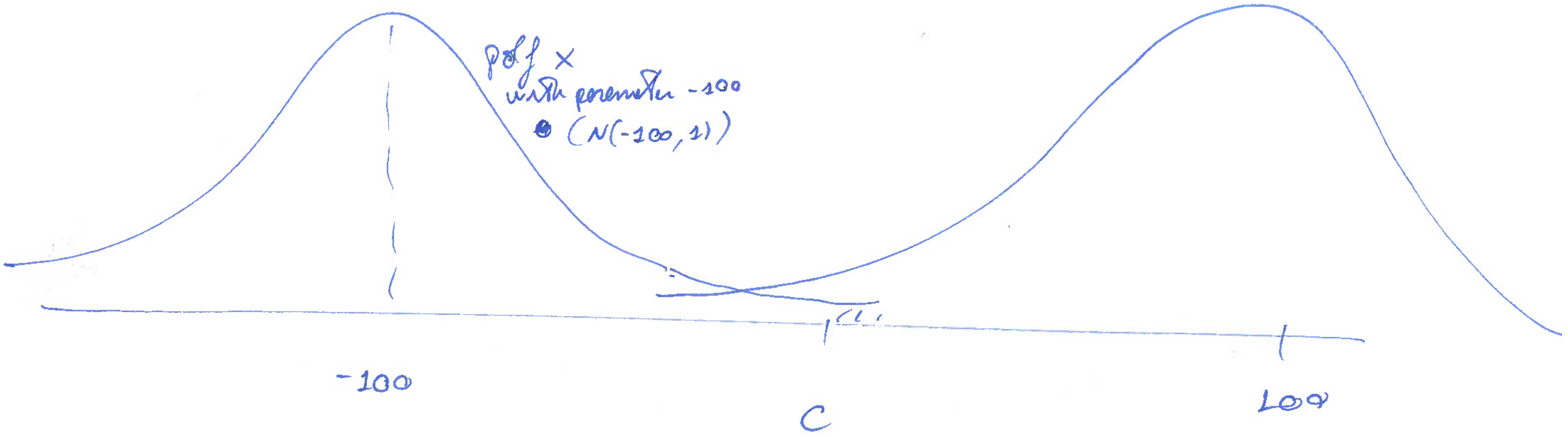
$$\beta(\theta) = P_\theta(\text{reject } H_0) = P_\theta(X \geq c)$$

$P_{\theta=0}(\text{reject } H_0) = \alpha$

Power of the test



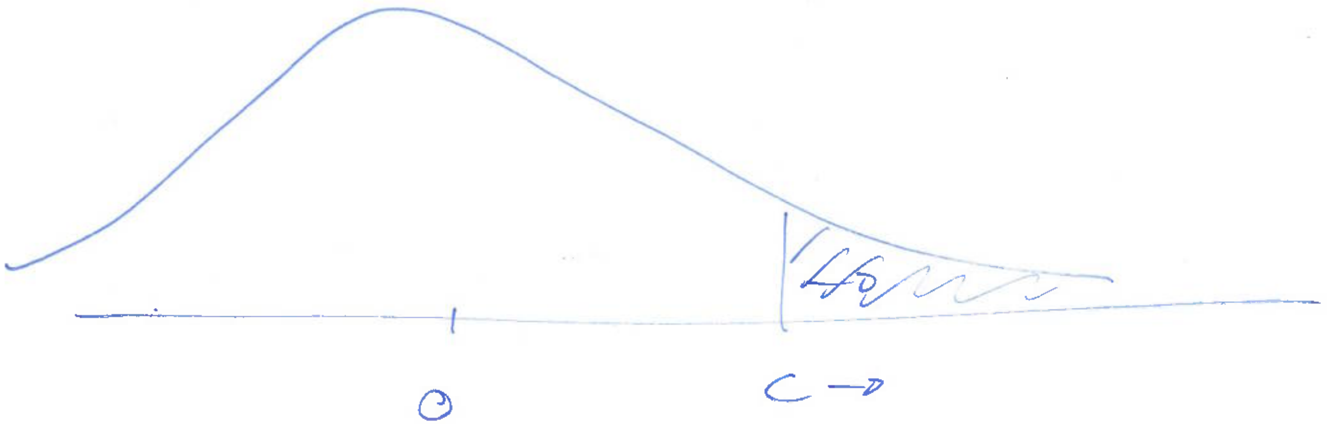




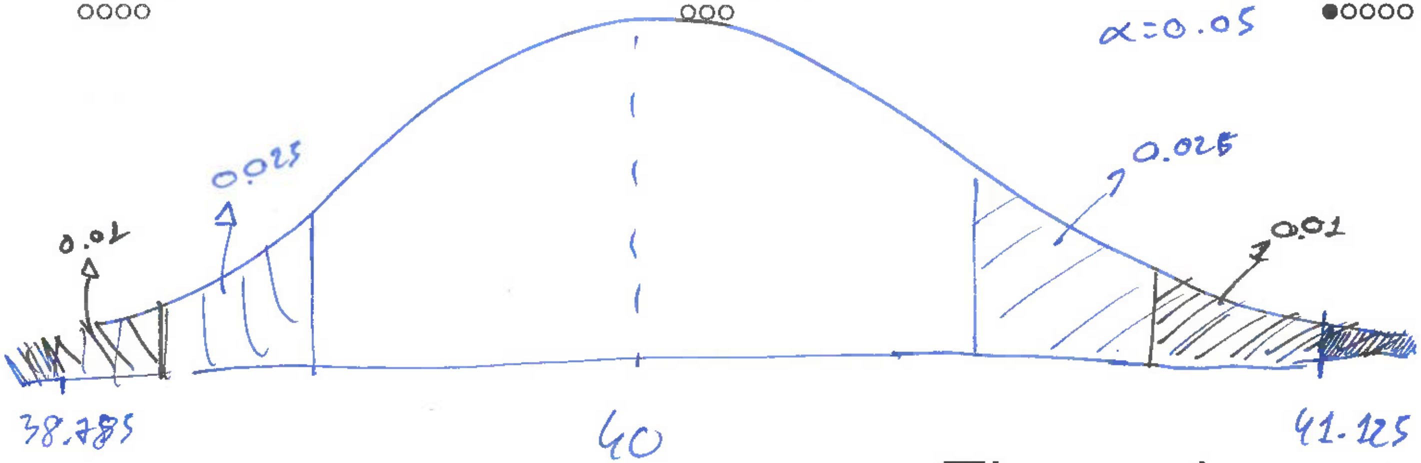
$P_{-100}(X \geq c)$ IS VERY SMALL

$P_{100}(X \geq c)$ IS VERY BIG

$\beta(\theta) = P_\theta(X \geq c)$



$\alpha = 0.05$ $\alpha = 0.02$



The p-value

WHAT IS THE LOWEST VALUE OF α FOR WHICH WE REJECT H_0 ?

$\Rightarrow$ α s.t. $\alpha = \cancel{P_{40}(|\bar{X}| \geq 41.125)}\ 2 \cdot P_{40}(\bar{X} \geq 41.125) = 1 - P_{40}(38.875 \leq \bar{X} \leq 41.125)$

Definition

Often the so-called p-value is reported (instead of a test decision):

$$p = \sup_{\theta \in \Theta_0} P_\theta(\text{observing something "at least as extreme" as the observed values})$$

Reject H_0 iff $p \leq \alpha \rightarrow \alpha$-level test.
If the test is based on the statistic T with rejection for large values of T then

$$p = \sup_{\theta \in \Theta_0} P_\theta(T \geq t),$$

where t is the observed value.
In the previous example (where $X \sim N(\theta, 1)$ and $H_0 : \theta \leq 0$ against $H_1 : \theta > 0$) the p-value is:

$$p = \sup_{\theta \in \Theta_0} P_\theta(X \geq x) = P_0(X \geq x) = 1 - \Phi(x)$$

Example: $X_1, \ldots, X_n \sim N(\mu, 1)$ iid, μ unknown

$$H_0 : \mu = \mu_0 \text{ against } H_1 : \mu \neq \mu_0$$

Level α test
Under H_0: $T = \sqrt{n}(\bar{X} - \mu_0) \sim N(0,1)$. Rejection region (based on T):

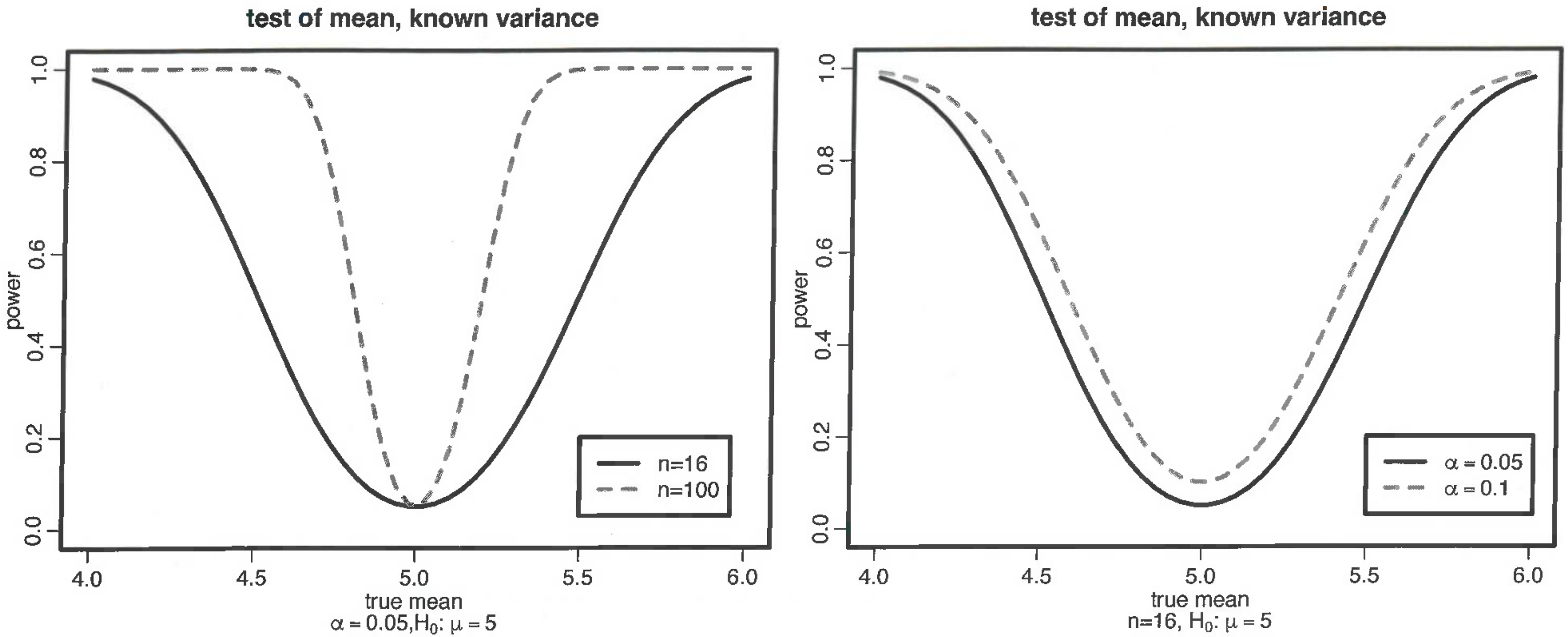
$$(-\infty, -c_{\alpha/2}] \cup [c_{\alpha/2}, \infty),$$

where $\Phi(c_{\alpha/2}) = 1 - \alpha/2$.
Test rejects for large values of $|T|$.
Hence, for the observation t the p-value is:

$$p = P_{\mu_0}(|T| \geq |t|) = P(T \leq -|t| \text{ or } T \geq |t|) = \Phi(-|t|) + 1 - \Phi(|t|) = 2 - 2\Phi(|t|)$$

Power: Note that $T \sim N(\sqrt{n}(\mu - \mu_0), 1)$.

$$\begin{aligned}\beta(\mu) =& P_\mu(|T| \geq c_{\alpha/2}) = 1 - P_\mu(-c_{\alpha/2} \leq T \leq c_{\alpha/2})\\ =& 1 - P_\mu(-\sqrt{n}(\mu - \mu_0) - c_{\alpha/2} \leq T - \sqrt{n}(\mu - \mu_0) \leq -\sqrt{n}(\mu - \mu_0) + c_{\alpha/2})\\ =& 1 - \Phi(-\sqrt{n}(\mu_0 - \mu) + c_{\alpha/2}) + \Phi(-\sqrt{n}(\mu_0 - \mu) - c_{\alpha/2})\end{aligned}$$



AS N INCREASES OUR ESTIMATES BECOMES MORE PRECISE, NAMELY OUR POWER FUNCTION BECOMES LARGER FOR $\theta \in \Theta_1 = \mathbb{R} \setminus \{5\}$, AND SO OUR TEST BECOMES MORE POWERFUL

Example: Student's t-Test; One-Sample t-Test

$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ iid, μ and σ unknown parameters

$$H_0 : \mu = \mu_0 \text{ against } H_1 : \mu \neq \mu_0$$

Under H_0: $T = \sqrt{n}\frac{\bar{X}-\mu_0}{S} \sim t_{n-1}$.
Rejection region:

$$(-\infty, -c] \cup [c, \infty),$$

where $c = t_{n-1,\alpha/2}$. ($t_{n-1,\alpha/2}$ is chosen such that if $Y \sim t_{n-1}$ then $P(Y > t_{n-1,\alpha/2}) = \alpha/2$)

Relating Tests and CIs

Constructing a test from a confidence region

Let Y be the random observations. Suppose $A(Y)$ is a $1-\alpha$ confidence region for θ, i.e.

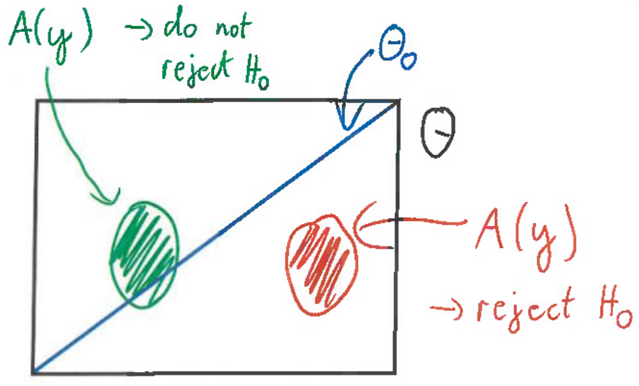
$$P_\theta(\theta \in A(Y)) \geq 1-\alpha \quad \forall \theta \in \Theta.$$

Then one can define a test for

$$H_0 : \theta \in \Theta_0 \quad \text{v.s.} \quad H_1 : \theta \notin \Theta_0$$

(where Θ_0 is some fixed subset of Θ) with level α as follows:

$$\text{Reject } H_0 \text{ if } \ \Theta_0 \cap A(y) = \emptyset.$$



To see that the above test has the appropriate level: For all $\theta \in \Theta_0$,

$$P_\theta(\text{type I error}) = P_\theta(\overset{H_0}{\text{reject}}) = P_\theta(\Theta_0 \cap A(Y) = \emptyset) \leq P_\theta(\theta \notin A(Y)) \leq \alpha.$$

Example: $Y \sim \text{Binomial}(n, \theta)$, $\theta \in (0, 1)$ unknown

$\sqrt{n}\frac{Y/n-\theta}{\sqrt{\theta(1-\theta)}}$ is approx. $N(0, 1)$ BY CLT

Leads to the confidence limits

$$\underbrace{\frac{y}{n} \pm \frac{c_{\alpha/2}}{\sqrt{n}}\sqrt{\frac{y}{n}(1-\frac{y}{n})}}$$

A level α test of $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ is defined by

$$\underline{R} = \left\{ y : \theta_0 \notin \overbrace{\left(\frac{y}{n} - \frac{c_{\alpha/2}}{\sqrt{n}}\sqrt{\frac{y}{n}(1-\frac{y}{n})}, \frac{y}{n} + \frac{c_{\alpha/2}}{\sqrt{n}}\sqrt{\frac{y}{n}(1-\frac{y}{n})} \right)}^{A(y)} \right\}$$

$y = Y(\omega)$

$P_{\theta_0}(\text{REJECT } H_0) = P_{\theta_0}(Y \in R) = P_{\theta_0}(\{\omega \in \Omega : Y(\omega) \in R\}) \underline{=} P_{\theta_0}(\{\omega \in \Omega : \theta_0 \notin A(Y(\omega))\})$

$= P_\theta(\theta_0 \notin A(Y)) \leq \alpha$

Next lecture

We consider how to use parametric likelihoods to derive further tests

Imperial College London

Lecture 10: Likelihood Ratio Tests

Statistical Modelling I

Dr. Riccardo Passeggeri

Outline

Introduction

Motivation

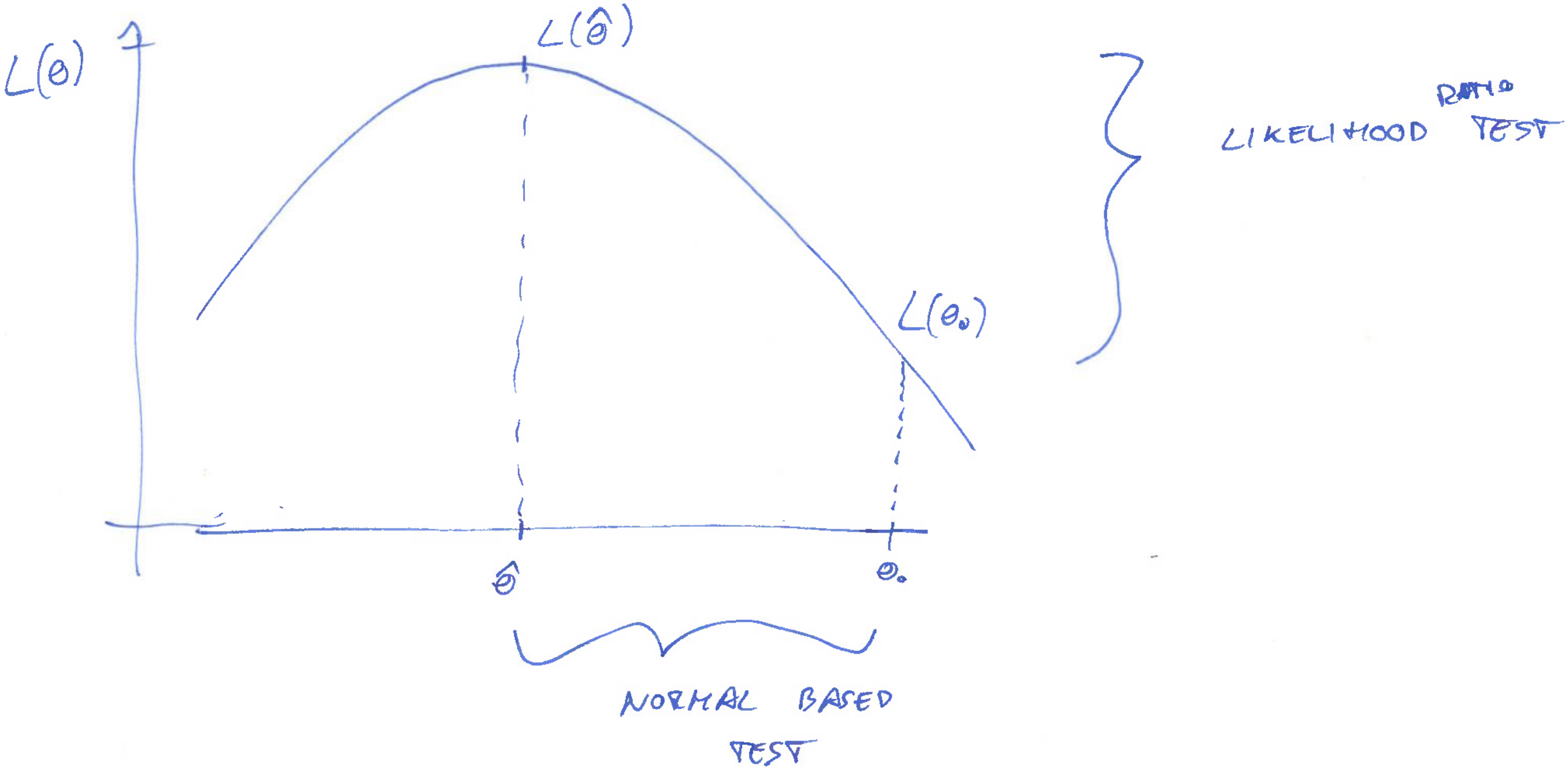
- Idea behind the maximum likelihood estimator: parameter with the highest likelihood is "best". Can this idea be used to create a test?
- More precisely, consider the hypotheses

$$H_0 : \theta \in \Theta_0 \quad \text{against} \quad H_1 : \theta \in \Theta_1 := \Theta \setminus \Theta_0$$

- Main idea: compare the maximised likelihood L under H_0 ($\sup_{\theta \in \Theta_0} L(\theta)$) to the unrestricted maximum likelihood ($\sup_{\theta \in \Theta} L(\theta)$). If the latter is (much?) larger then $\sup_{\theta \in \Theta_1} L(\theta) >> \sup_{\theta \in \Theta_0} L(\theta)$, casting doubt on H_0.

Example: $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

For MLEs, we have seen the (approx.) pivotal quantity $\sqrt{nI_f(\hat{\theta})}(\hat{\theta} - \theta_0)$ $\dot\sim N(0,1)$



Definition: Likelihood Ratio Test Statistic

[redacted]

Suppose we observe the data y. The **likelihood ratio test statistic** is

$$t(y) = \frac{\sup_{\theta \in \Theta} L(\theta; y)}{\sup_{\theta \in \Theta_0} L(\theta; y)} \overset{= L(\hat{\theta}_{MLE})}{=} \frac{\text{max. lik. under } H_0 + H_1}{\text{max. lik. under } H_0}$$

- If $t(y)$ is "large" this will indicate support for H_1, so reject H_0 when

$$t(y) \geq k,$$

where k is chosen to make

$$\sup_{\theta \in \Theta_0} P_\theta(t(Y) \geq k) = (\text{or} \leq) \alpha$$

(e.g. $\alpha = 0.05$).

- The choice of k ensures that we get a test to the level α.

Examples

Example 1: $X \sim$Binomial(n, θ), $\theta \in (0,1) = \Theta$

$$H_0 : \theta = 0.5 \quad \text{v.s.} \quad H_1 : \theta \neq 0.5$$

- Here, $\Theta_0 = \{0.5\}$, $\Theta_1 = (0, 0.5) \cup (0.5, 1)$
- The likelihood is

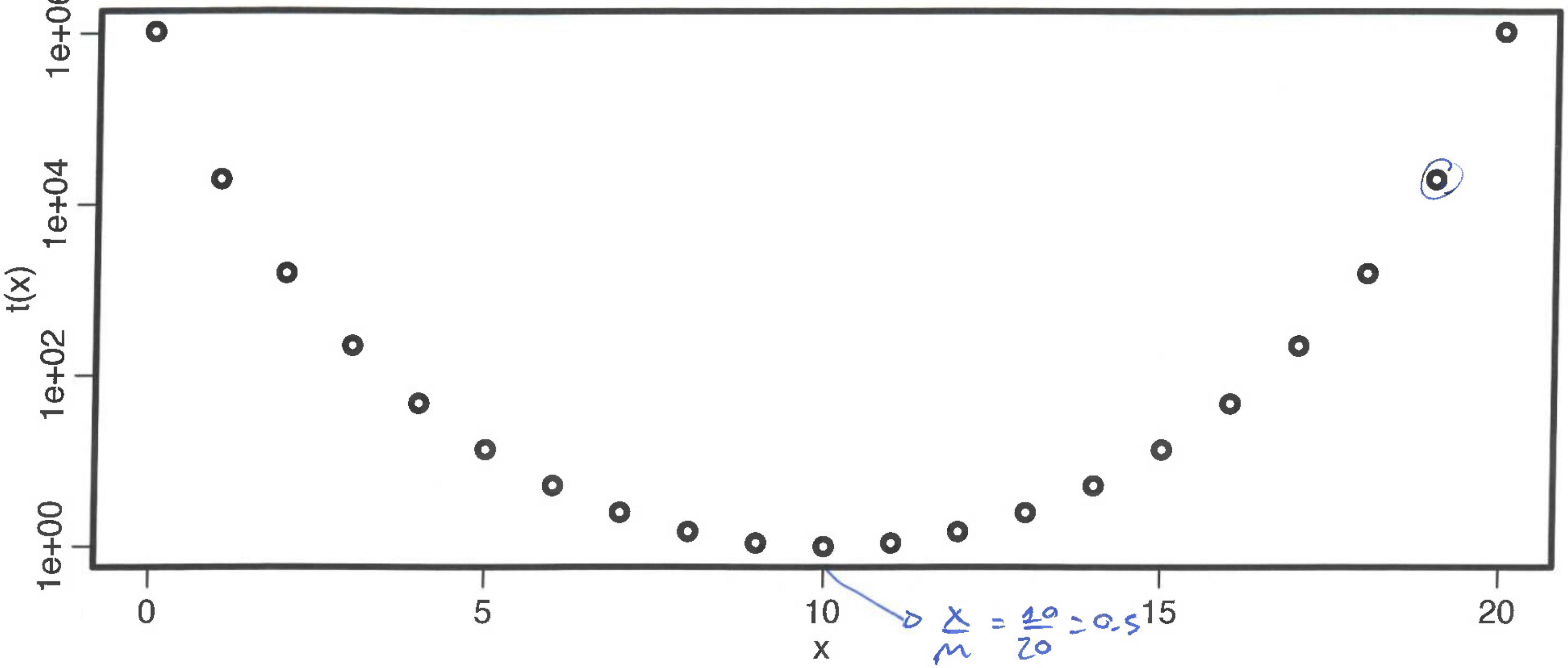
$$L : \Theta \to \mathbb{R}, \theta \mapsto P_\theta(X = x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}.$$

- The LRT statistic is

MLE

$$t(x) = \frac{\sup_{\theta \in \Theta} L(\theta)}{\sup_{\theta \in \Theta_0} L(\theta)} = \frac{L(\frac{x}{n})}{L(0.5)} = \frac{\left(\frac{x}{n}\right)^x \left(1 - \frac{x}{n}\right)^{n-x}}{0.5^n}$$

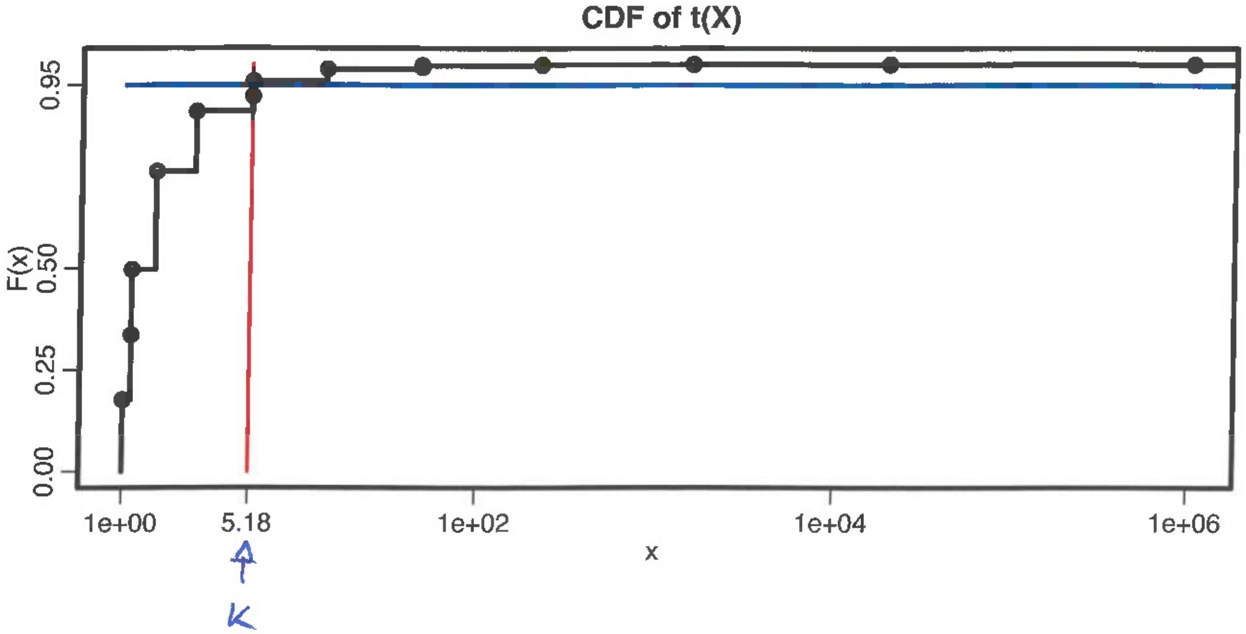
Dependence of the likelihood ratio t on the observation x.



$\frac{x}{n} = \frac{10}{20} = 0.5$

The above is for n=20. Note the log–scale on the y–axis.

- To construct the test we need the distribution of t under H_0



$P_{\theta_0}(t(X) \geq k) \leq \alpha = 0.05$

$P_{\theta_0}(t(x) \geq 5.18) \leq 0.05$

- In this case rejecting if $t > 5.19$ leads to a test with level 5%

Example 2: $X_i \sim \text{Binomial}(n, \theta_i), i = 1, 2$ independent

$$H_0 : \theta_1 = \theta_2 \text{ v.s. } H_1 : \theta_1 \neq \theta_2$$

$\left(\frac{x_1}{n}, \frac{x_2}{n}\right)$

The LRT statistic is

$$t = \frac{L(x_1/n, x_2/n)}{L(\frac{x_1+x_2}{2n}, \frac{x_1+x_2}{2n})}$$

The distribution of t under H_0 is not easy to obtain as for each value of $(\theta_1, \theta_2) \in \Theta_0$ the distribution of t may be different.

Example 3: m factories producing light bulbs. Are all factories producing bulbs of the same quality?

Observation: life-length of n light bulbs from each factory; Y_{ij} = life-length of bulb j from factory i.

Model: Y_{ij} indep. Exp(λ_i), i=1,...,m; j=1,...n, $\lambda_i > 0$ unknown, $i = 1, \ldots, m$.

$$H_0 : \lambda_1 = \cdots = \lambda_m \quad \text{v.s.} \quad H_1 : \text{ not } H_0$$

Under $H_0 + H_1$ (using $\theta^t = (\lambda_1, \dots, \lambda_m)$):

$$L(\theta) = \prod_{i=1}^{m} \prod_{j=1}^{n} \lambda_i \exp(-\lambda_i Y_{ij}) = \prod_{i=1}^{m} \lambda_i^n e^{-\lambda_i \sum_j y_{ij}}$$

$\implies$ m likelihoods of iid Exponential(λ_i) observations, leading to the MLE

$$\hat{\lambda}_i = \frac{1}{\bar{y}_i} \text{ where } \bar{y}_i = \frac{1}{n} \sum_j y_{ij}.$$

Under H_0 (setting $\lambda := \lambda_1 = \cdots = \lambda_m$):

$$L(\theta; y) = \lambda^{mn} e^{-\lambda \sum_{i,j} y_{ij}}$$

$\implies$ the likelihood for iid Exponential(λ) observations
MLE is

$$\hat{\lambda} = \frac{1}{\bar{y}}$$

LRT statistic:

$$t(y) = \frac{\sup_{\theta \in \Theta} L(\theta; y)}{\sup_{H_0} L(\theta; y)} = \frac{\frac{e^{-mn}}{(\prod_i \bar{y}_i)^n}}{\frac{e^{-mn}}{\bar{y}^{mn}}} = \frac{\bar{y}^{mn}}{(\prod_i \bar{y}_i)^n}$$

Remarks

- To construct a test we would need to know the distr. of $t(Y)$ under H_0. (Not easy)
- Even if it were known - the distribution of $t(Y)$ may depend on λ and hence, choosing k according to $\sup_{\lambda>0} P_\lambda(t(Y) \geq k) = \alpha$ may not be easy.