STAT 346/446 Lecture 11

Most powerful tests and *p*-values

CB Sections 8.3.2 and 8.3.4, DS Sections 9.2 and 9.3

1. Most powerful tests
   - Neyman-Pearson Lemma (simple hypotheses)
   - Karlin-Rubin Theorem (composite hypotheses)
   - Most powerful unbiased tests
2. p-values

Note: We skip Section 8.3.3 for now

Power of a test

- We are looking for a good way to test for hypotheses $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_o^c$

  → Rejecting a true null hypothesis

- Want probability of Type I Error to be small
- For many methods (e.g. LRT) we can control the probability of Type I Error
  - A level α test has Type I Error probabilities at most α for all $\theta \in \Theta_0$
- Also want probability of Type II Error to be small (high power)
  - power of a test = 1 - prob. of Type II Error
- In the class of *level α tests*, can we find the one i.e. the test with the smallest Type II Error probabilities?
  - I.e. the level α test with *maximum power* for all $\theta \in \Theta_0^c$?

Hypotheses: $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_0^c$

Test procedure (decision rule): Reject H_0 iff $\underline{x} \in R$

obs. sample

rejection region

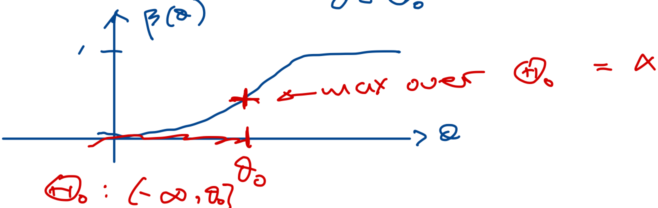
Power function:

$$\beta(\theta) = P_\theta(\text{Reject } H_0) = P_\theta(\underline{X} \in R)$$

Want: $\beta(\theta)$ to be small for $\theta \in \Theta_0$

$\beta(\theta)$ to be large for $\theta \in \Theta_0^c$

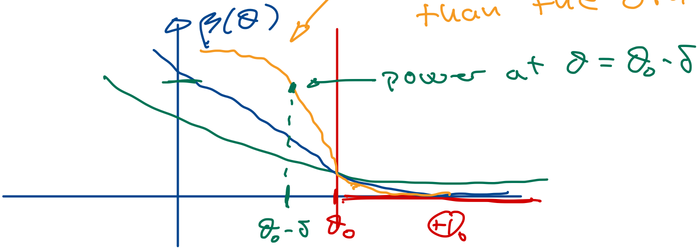
Size of a test: $\alpha = \sup_{\theta \in \Theta_0} \beta(\theta)$



Most powerful test

~~2~~ 3 size α tests:

uniformly more powerful than the other two



Two test procedures:

Test δ: Reject H_0 iff. $X \in R$

Test δ^*: Reject H_0 iff. $X \in R^*$

Power of a test - sample size calculations

- In applied stats courses we often talk about the power of a test at a particular θ value in Θ_0^c

Example

- Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\mu, \sigma^2)$
- The level α z-test for $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$ rejects H_0 if

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha$$

- Can calculate the power for $\mu = \mu_0 + \delta$

$$\beta(\mu_0 + \delta) = P(Z_0 > z_\alpha | \, \mu = \mu_0 + \delta) = 1 - \Phi\left(z_\alpha - \delta\sqrt{n}/\sigma\right)$$

and can choose n to get desired power at a particular value of $\mu \in (\mu_0, \infty)$

Most powerful tests

Def: Uniformly most powerful

Let $\mathcal{C}$ be a class of tests for testing $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_0^c$.

A test $\delta \in \mathcal{C}$ is a **uniformly most powerful (UMP) class $\mathcal{C}$ test** if for any other $\delta^* \in \mathcal{C}$ we have

$$\beta_\delta(\theta) \geq \beta_{\delta^*}(\theta) \qquad \forall \theta \in \Theta_0^c$$

alternative

- Focus on the class $\mathcal{C}$ = all level α tests.
- UMP is a very strong requirement and does not exists in many realistic problems.

UMP overview

- **Neyman-Pearson Lemma** - strong result for a limited case
  - Gives a level α UMP test for *simple hypotheses*

$$H_0 : \theta = \theta_0 \qquad \text{versus} \qquad H_1 : \theta = \theta_1$$

$\Theta_0 = \{\theta_0\}$
$\Theta_0^c = \{\theta_1\}$
i.e.
$\Theta = \{\theta_0, \theta_1\}$

  - Gives both necessary and sufficient conditions
  - Can be used to prove results for *composite hypotheses*

- **Karlin-Rubin Theorem** - strong result for another limited case
  - Gives a level α UMP test for *one-sided hypotheses*

$$H_0 : \theta \leq \theta_0 \qquad \text{versus} \qquad H_1 : \theta > \theta_0$$

  - Requires a *monotone likelihood ratio* family for the test statistic
  - More practical than the Neyman-Pearson

Theorem: Neyman-Pearson Lemma

Consider hypotheses $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$ and a test procedure with rejection region R where

$R \cup R^c = \mathcal{X}$ = sample space

1. for some $k \geq 0$ we have

description of R

$$\mathbf{x} \in R \text{ if } f(\mathbf{x}|\theta_1) > k \overbrace{f(\mathbf{x}|\theta_0)}^{\text{likelihood.}}$$

$$\text{and } \mathbf{x} \in R^c \text{ if } f(\mathbf{x}|\theta_1) < k\, f(\mathbf{x}|\theta_0)$$

$\Rightarrow$ Reject if $\dfrac{f(\underline{x}|\theta_1)}{f(\underline{x}|\theta_0)} > k$ for some $k \geq 0$

2. and $\alpha = P_{\theta_0}(\mathbf{X} \in R)$ $= \beta(\theta_0)$

Then

a) (Sufficiency) Any test that satisfies 1 and 2 is a UMP level α test

b) (Necessity) If there exists a test that satisfies 1 and 2 with $k > 0$ then
   - every UMP level α test is a size α test, and
   - every UMP level α satisfies 1 except perhaps on a set A where $P_{\theta_0}(\mathbf{X} \in A) = P_{\theta_1}(\mathbf{X} \in A) = 0$

Neyman-Pearson Lemma

- **Simple** hypothesis: Θ_0 (or Θ_0^c) contains exactly one point
- **Composite** hypothesis: Θ_0 (or Θ_0^c) contains more than one point
- To use the Neyman-Pearson Lemma we proceed as follows:
  - Find the joint distribution of $X_1, X_2, \ldots, X_n$
  - Express the ratio $\frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)} > k$ in terms of a statistic
  - Choose k such that it is a level α test.

  then by Neyman-Pearson we have a UMP level α test
- By factorization theorem $\frac{f(\mathbf{x}|\theta_1)}{f(\mathbf{x}|\theta_0)}$ should be a function of a sufficient statistics.

Neyman-Pearson for sufficient statistic

Corollary

Consider hypotheses $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$.

Suppose T is a sufficient statistic for θ with pdf or pmf $g(t|\theta_i)$ corresponding to θ_i, $i = 0, 1$.

Then any test based on T with rejection region S is a UMP level α test if it satisfies

1. for some $k \geq 0$ we have

Reject if $\frac{g(t|\theta_1)}{g(t|\theta_0)} > k$

$$t \in S \text{ if } g(t|\theta_1) > k\, g(t|\theta_0)$$
$$\text{and } t \in S^c \text{ if } g(t|\theta_1) < k\, g(t|\theta_0)$$

2. and $\alpha = P_{\theta_0}(T \in S)$

$\longrightarrow$ Can work with pdf of the test statistic instead of the joint pdf.

Notes on the NP lemma

Connection with the LRT

- Neyman-Pearson: Reject if

actually a LRT, just written differently

$$\frac{f(\mathbf{x} \mid \theta_1)}{f(\mathbf{x} \mid \theta_0)} > k \qquad \text{for some } k \geq 0$$

Note that θ_0 is in the denominator

- The LRT rejects if

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \theta_0} L(\theta \mid \mathbf{x})}{\sup_{\theta \in \theta} L(\theta \mid \mathbf{x})} \leq c \quad \text{for some } c \in [0, 1]$$

if $\Theta_0 = \{\theta_0\}$, $\Theta_0^c = \{\theta_1\}$

$$= \frac{f(\mathbf{x} \mid \theta_0)}{\max\{f(\mathbf{x} \mid \theta_0), f(\mathbf{x} \mid \theta_1)\}} = \begin{cases} 1 & \text{if } f(\underline{x} \mid \theta_0) > f(\underline{x} \mid \theta_1) \\ \dfrac{f(\underline{x} \mid \theta_0)}{f(\underline{x} \mid \theta_1)} & \text{o.w.} \end{cases}$$

Here θ_0 is in the numerator

Notes on the NP lemma

Role of k

- Neyman-Pearson: Reject if

$$f(\mathbf{x} \mid \theta_1) > k\, f(\mathbf{x} \mid \theta_0) \qquad \text{for some } k \geq 0$$

- Why not just pick the θ value that gives the larger likelihood?
    - That is, reject if $f(\mathbf{x} \mid \theta_1) > f(\mathbf{x} \mid \theta_0)$ ($k = 1$)
- Reason:
    - Want to control $\beta(\theta_0) =$ probability of Type I error
    - α determines the k

Example

- Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\theta, 1)$
- Find the size α UMP for testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$

Example

- Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{Poisson}(\theta)$
- Find the size α UMP for testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1$

Monotone likelihood ratio families

- Plan: Use Neyman-Pearson Lemma to get results for composite hypotheses
- First: define monotone likelihood ratio

Def: MLR

A family of pdfs/pmfs $\{g(t|\theta) : \theta \in \Theta\}$ for a univariate random variable T and real-valued parameter θ has a **monotone likelihood ratio (MLR)** if for every $\theta_2 > \theta_1$ the ratio

as a function of t

$$LR(t) = \frac{g(t|\theta_2)}{g(t|\theta_1)}$$

i.e on the support of g

is a monotone function of t on $\{t : g(t|\theta_1) > 0 \text{ or } g(t|\theta_2) > 0\}$

- Many common families of distributions have an MLR

Example

- Show that the normal family and the Poisson family have an MLR

Normal: Let $T \sim N(\theta, \sigma^2)$, σ^2 known, $\theta_2 > \theta_1$

$$LR(t) = \frac{g(t|\theta_2)}{g(t|\theta_1)} = \frac{\frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(t-\theta_2)^2}{2\sigma^2}\right)}{\frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(t-\theta_1)^2}{2\sigma^2}\right)}$$

$$= \exp\left(-\frac{1}{2\sigma^2}\left(t^2 - 2t\theta_2 + \theta_2^2 - t^2 + 2t\theta_1 - \theta_1^2\right)\right)$$

$$= \underbrace{\exp\left(-\frac{1}{2\sigma^2}(\theta_2^2 - \theta_1^2)\right)}_{\text{no } t \text{ in here}} \exp\left(\frac{t\overbrace{(\theta_2 - \theta_1)}^{>0}}{\sigma^2}\right)$$

e^x is a monotone increasing function of x

$\Rightarrow$ as t increases, $\frac{t(\theta_2-\theta_1)}{\sigma^2}$ increases and $LR(t)$ increases

Poisson: $T \sim \text{Poisson}(\theta)$, $\theta_2 > \theta_1$

$$LR(t) = \frac{g(t|\theta_2)}{g(t|\theta_1)} = \frac{e^{-\theta_2}\,\theta_2^{\,t}\,\frac{1}{t!}}{e^{-\theta_1}\,\theta_1^{\,t}\,\frac{1}{t!}}$$

$$= \underbrace{e^{\theta_1-\theta_2}}_{\text{const. as as a funct. of } t} \left(\frac{\theta_2}{\theta_1}\right)^t \qquad \frac{\theta_2}{\theta_1} > 1 \implies \left(\frac{\theta_2}{\theta_1}\right)^t \text{ increases as } t \text{ increases.}$$

$\Rightarrow$ $LR(t)$ is a monotone increasing function of t

$\Rightarrow$ The Poisson family has a MLR.

Karlin-Rubin Theorem

Theorem: Karlin-Rubin

Consider testing $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$

Suppose T is a sufficient statistic for θ and the family of pdfs/pmfs of T has an MLR.

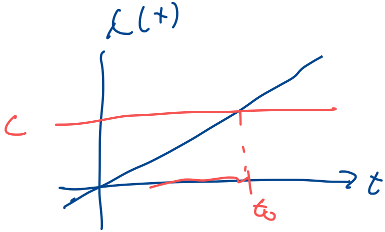
Then for any t_0, the test that rejects H_0 if and only if $T > t_0$ is a UMP level α test, where $\alpha = P_{\theta_0}(T > t_0)$ ← use this to find t_0

Similarly:

- If testing $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$
- then a test that rejects H_0 if and only if $T < t_0$ is a UMP level α test.

e.g. LRT : Reject

$$\lambda(t) = \frac{L(\theta_0 | t)}{L(\theta_1 | t)} \le c$$



$\lambda(t) < c$

$\Rightarrow \quad t < t_0$

Power function

Setup: $H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_0^c$

Reject H_0 if $T(\underline{X}) \in R$

↳ some test statistic

e.g. $\lambda(t) \le c$

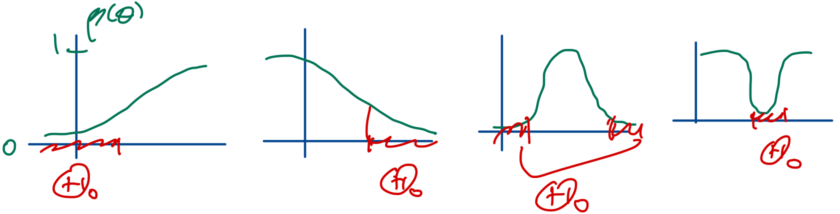
Power function:

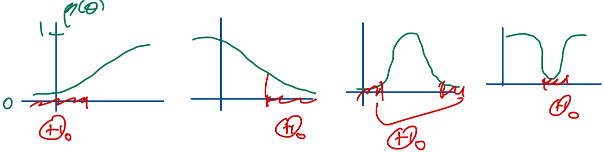
$$\beta(\theta) = P_\theta(\text{Reject } H_0) = P_\theta(T(\underline{X}) \in R)$$

→ Need the distribution of T to evaluate $\beta(\theta)$

→ Function of θ e.g.

want $\beta(\theta)$ small in Θ_0 and large in Θ_0^c





size, level, most powerful, prob. of Type I & II errors are characteristics of $\beta(\theta)$ over either Θ_0 or Θ_0^c <u>only</u>

e.g. size: $\alpha = \sup_{\theta \in \Theta_0} \beta(\theta)$

$\beta(\theta)$ = prob of type I error for $\theta \in \Theta_0$

$\beta(\theta) = 1 -$ prob of type II error for $\theta \in \Theta_0^c$

Test is uniformly most powerful if

$$\beta(\theta) \geq \beta_*(\theta) \quad \forall \theta \in \Theta_0^c$$

for any other test $*$

Idea behind proof of Karlin-Rubin

NP

$\frac{L(\theta_1 | x)}{L(\theta_0 | x)}$

+ MLR

- Take any $\theta' \in \Theta_0^c = (\theta_0, \infty)$
- Use Neyman-Pearson lemma to show that the test:

$$\text{Reject } H_0 \text{ iff } T > t_0 \text{ for some } t_0$$

any $\theta' \in \Theta_0^c$

is a UMP level α test for $H_0 : \theta = \theta_0 \qquad \text{vs.} \qquad H_1 : \theta = \theta'$

  - Need the MLR property to show this

- Since this holds for any $\theta' \in \Theta_0^c$ (and because of MLR) it follows that this test is the UMP level α test for

i.e. any θ in Θ_0^c

$$H_0 : \theta = \theta_0 \qquad \text{vs.} \qquad H_1 : \theta > \theta_0$$

- Finally, it follows (MLR) that this test is the UMP level α test for

$$H_0 : \theta \leq \theta_0 \qquad \text{vs.} \qquad H_1 : \theta > \theta_0$$

Example - Normal

MLR : property of the distribution of T

- Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{N}(\theta, \sigma^2)$, σ^2 known
- Find the size α UMP for testing $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$

Know: $\bar{X}$ is a suff. statistic for μ.

$\bar{X} \sim N(\cancel{\mu}\ \theta, \sigma^2/n)$

and the normal family has a MLR (showed last time)

$\Rightarrow$ By Karlin-Rubin the UMP level α test rejects H_0 if $\bar{X} > t_0$ where

$$\alpha = P_{\theta_0}(\bar{X} > t_0) = 1 - P\left(\frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}} \leq \frac{t_0 - \theta_0}{\sigma/\sqrt{n}}\right)$$

$$= 1 - \Phi\left(\frac{t_0 - \theta_0}{\sigma/\sqrt{n}}\right) \Rightarrow t_0 = \theta_0 + z_\alpha \sigma/\sqrt{n}$$

$t_0 = \theta_0 + z_\alpha \sigma/\sqrt{n}$

so the UMP level α test rejects H_0 if

$$\bar{X} > \theta_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$$

$$\iff \quad \frac{\bar{X} - \theta_0}{\sigma/\sqrt{n}} > z_\alpha$$

i.e. the usual z-test.

Example - Poisson

- Let $X_1, \ldots, X_n$ be a random sample from $\mathrm{Poisson}(\theta)$
- Find the size α UMP for testing $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$

(1) sufficient statistic: both $\bar{X}$ and $T = \sum_{i=1}^{n} X_i$ are sufficient statistics for θ

(2) know: $T \sim \text{Poisson}(n\theta)$

(3) and Poisson has a MLR $\Rightarrow$ Conditions of Karlin-Rubin are met

$\Rightarrow$ By Karlin-Rubin the UMP level α test rejects H_0 if $T < t_0$ where

$$\alpha = P_{\theta_0}(T < t_0) = \sum_{k=0}^{t_0 - 1} \frac{e^{-n\theta_0} (n\theta_0)^k}{k!}$$

use trial and error for a given θ_0 and n to find t_0

T is always a whole number so only need to consider $t_0 \in \mathbb{Z}$

Example - Poisson, continued

For $\theta_0 = 2.3$ and $n = 10$ Poisson cdf $t_0 - 1$ $n\theta_0$

```
t0 <- 9:18
n <- 10; theta0 <- 2.3
kable(cbind(t0, alpha = ppois(t0, lambda = theta0*n)))
```

$\alpha = 0.05$

$t_0 - 1$ ~~t0~~	alpha
9	0.0008060
10	0.0019775
11	0.0044270
12	0.0091219
13	0.0174282
14	0.0310743
15	0.0519983
16	0.0820766
17	0.1227707
18	0.1747687

$t_0 - 1 = 14 \Leftarrow$ size: 0.031

size: 0.052

$\Rightarrow t_0 = 15$

$$\alpha = P_{\theta_0}(T < t_0) = \sum_{k=0}^{t_0-1} \frac{e^{-n\theta_0} (n\theta_0)^k}{k!}$$

take $t_0 - 1$ that gives the largest α value less than 0.05

Shrinking the class of tests

Def: Unbiased tests

A test with power function $\beta(\theta)$ is **unbiased** if

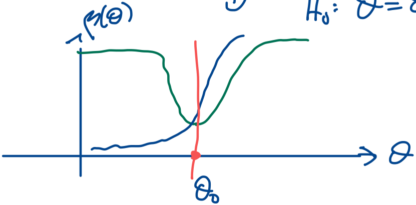
$$\beta(\theta_0) \geq \beta(\theta_1) \qquad \text{for all } \theta_0 \in \Theta_0, \theta_1 \in \Theta_0^c$$

- UMP level α tests usually don't exists for two-sided hypotheses $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

more powerful on (θ_0, ∞)
but not on $(-\infty, \theta_0)$

- But sometime a UMP unbiased level α test exists (example 8.3.20 in the book)

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



$\Theta_0 = \{\theta_0\}$

p-value

have chosen the level of the test

- Reporting conclusions from a hypothesis test:
  - One way: H_0 is rejected (or not rejected) with significance level α.
  - Another way: Report a p-value

The "user" decides the level of the test.

Definition: p-value

(1)
- A **p-value $p(\mathbf{X})$** is a test statistic that satisfies i.e. a statistic

$$0 \le p(\mathbf{x}) \le 1 \qquad \forall \mathbf{x} \in \mathcal{X}$$

(2)
- A p-value is called a **valid p-value** if

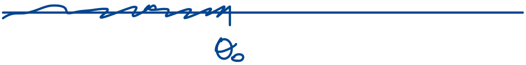
$$P_\theta(p(\mathbf{X}) \le \alpha) \le \alpha \qquad \forall \theta \in \Theta_0,\ \forall \alpha \in [0, 1]$$

$$P_\theta(p(\mathbf{X}) \leq \alpha) \leq \alpha \qquad \forall \theta \in \Theta_0,\ \forall \alpha \in [0, 1]$$

Don't want to reject θ

θ in here i.e. H_0 is true

Prob(p-value $\leq \alpha$) $\leq \alpha$



$\Theta_0 = (-\infty, \theta_0]$

Constructing a level α test

- If we have a valid *p*-value $p(\mathbf{X})$, we can use it to construct a level α test:

Reject H_0 iff $p(\underline{X}) \leq \alpha$

This is a level α test since

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta \in \Theta_0} P_\theta(\text{Reject } H_0)$$

$$= \sup_{\theta \in \Theta_0} P_\theta(p(\underline{X}) \leq \alpha) \overset{*}{\leq} \alpha$$ since $p(\underline{X})$ is a valid p-value

* holds for all $\theta \in \Theta_0$ so also for the maximizer

$\Rightarrow$ Have a level α test.

Finding a valid p-value

Theorem 8.3.27

Let $W(\mathbf{X})$ be a test statistic such that large values of W give evidence that H_1 is true. For each sample point $\mathbf{x}$ set

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} P_\theta\left(W(\mathbf{X}) \geq W(\mathbf{x})\right)$$

Then $p(\mathbf{X})$ is a valid p-value

- "Large values of W give evidence that H_1 is true" $\Rightarrow$ Reject H_0 for large values of $W(\underline{X})$

  $\Rightarrow$ Reject H_0 if $W(\underline{X}) > c$ for some c

  (c depends on α level)

Reject if $W(\underline{x}) > c$

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} P_\theta\big(W(\mathbf{X}) \geq W(\mathbf{x})\big)$$

$W(\underline{x})$ = observed test statistic

lower case $\underline{x}$

$W(\underline{X})$ is a random variable

upper case $\underline{X}$

maximum prob. over all $\theta \in \Theta_0$ (H_0 is true) that we observe something more extreme ("$\geq$") than our sample $\underline{x}$.

i.e. P-value: Prob. that we can get something more extreme than we actually got, given that H_0 is true.

Here, this means $\sup_{\theta \in \Theta_0}$

$$p(\underline{x}) = \sup_{\theta \in \Theta_0} P_\theta \left(W(\underline{X}) \geq W(\underline{x}) \right)$$

Lets define $w = W(\underline{x})$

lower case

$$p(w) = \sup_{\theta \in \Theta_0} P\left(W(\underline{X}) \geq w \right)$$

← function of $w = W(\underline{x})$ i.e. function of data.

$$\Rightarrow p(W) = \sup_{\theta \in \Theta_0} P\left(W(\underline{X}) \geq W \right)$$

upper case

is a random variable

Example

$\mu \le \mu_0$ $\quad$ σ^2 anything

$\Theta_0 : (-\infty, \mu_0] \times (0, \infty)$

- $X_1, \ldots, X_n$ iid $N(\theta, \sigma^2)$. Find valid p-values for the two-sided and one-sided t-test.

Two-sided: $\quad H_0 : \mu \le \mu_0 \quad$ vs $\quad \mu > \mu_0$

LRT: Reject for larg values of $T = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = W(X)$

Recall: $\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$ if μ is the true mean

A valid p-value:

$$p(\underline{x}) = \sup_{(\mu,\sigma^2) \in \Theta_0} P(W(\underline{X}) \ge W(\underline{x}))$$

$$= \sup_{\Theta_0} P\left(\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \ge W(\underline{x}) \right)$$

A valid p-value:

$\mu - \mu_0$ negative

$$p(\underline{x}) = \sup_{(\mu,\sigma^2)\in\Theta_0} P(W(\underline{X}) \geq W(\underline{x}))$$

$$= \sup_{\Theta_0} P\left(\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \geq W(\underline{x})\right)$$

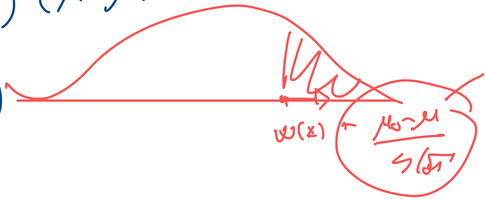
$$= \sup_{\Theta_0} P\left(\underbrace{\frac{\bar{X}-\mu}{S/\sqrt{n}}}_{\sim\, t_{n-1}} \geq W(x) \underbrace{- \frac{\mu-\mu_0}{S/\sqrt{n}}}_{\text{positive}}\right)$$

$\mu \in (-\infty, \mu_0]$

$$= P(T_{n-1} \geq w(\underline{x})) \quad \text{setting } (\mu = \mu_0)$$

$$= P\left(T_{n-1} \geq \frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right)$$

t_{n-1}



largest at $\mu = \mu_0$

Two-sided: Reject if

$$W(\underline{x}) = \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right| > c$$

A valid p-value:

$$p(x) = \sup_{\substack{(\mu, \sigma^2) \\ \in \Theta_0}} P\left(W(\underline{X}) \geq W(\underline{x}) \right)$$

$$\Theta_0 = \{\mu_0\} \times (0, \infty)$$

Just one point

$$= P_{\mu_0}\left(W(X) \geq \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right)$$

$W(X) \sim t_{n-1}$