Hypothesis testing framework

$H_0: \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\leftarrow$ χ^2_2 $\leftarrow$ length of part of β that we're testing

Last time

We have the model

$$Y_i \sim Bernoulli(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 SecondClass_i + \beta_4 ThirdClass_i$$

We want to test whether there is a difference in the chance of survival for second and third class passengers, holding age and sex fixed.

> What hypotheses should we test?

$H_0: \beta_3 = \beta_4$ $\qquad$ $H_A: \beta_3 \neq \beta_4$

$\beta_4 - \beta_3 = 0$ $\qquad$ $\beta_4 - \beta_3 \neq 0$

want: C st $C\beta = \beta_4 - \beta_3$

Contrasts

$$\beta_4 - \beta_3 = \underbrace{[0\ 0\ 0\ \ -1\ \ 1]}_{a^T} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_4 \end{bmatrix} = a^T \beta$$

$= C\beta$

$C = a^T$

$$\boxed{\begin{array}{l} \underline{\text{Def}}: \text{ Let } \theta \in \mathbb{R}^q \text{ and } a \in \mathbb{R}^q \text{ such that } \sum_i a_i = 0 \\ \text{Then } a^T\theta \text{ is called } \underline{\text{contrast}} \end{array}}$$

$H_0: a^T\theta = \gamma_0$ $\quad$ $H_A: a^T\theta \neq \gamma_0$

$\underline{\text{ex}}$: $H_0: a^T\beta = 0$ $\quad$ $H_A: a^T\beta \neq 0$

$\underline{\text{Test}}$: $Z_n = \left(a^T I^{-1}(\beta) a\right)^{-\frac{1}{2}} (a^T\hat{\beta} - 0) = \dfrac{a^T\hat{\beta}}{\sqrt{a^T I^{-1}(\beta) a}}$

$\approx N(0,1)$ under H_0

reject H_0 when $|Z_n| > z_{\frac{\alpha}{2}}$ (or $Z_n^2 > \chi^2_{1,\alpha}$)

Class activity

Work on Part II from last class:

https://sta711-s23.github.io/class_activities/ca_lecture_18.html

Class activity

$\mathcal{I}^{-1}(\beta)$

$\hat{\beta}$

```
a <- c(0, 0, 0, -1, 1)
test_stat <- (t(a) %*% coef(m1))/sqrt(t(a) %*% vcov(m1) %*% a)
test_stat
```

$a^T\hat{\beta}$ $\qquad$ $\sqrt{a^T \mathcal{I}^{-1}(\beta) a}$

```
##              [,1]
## [1,] -5.207289
```

$Z_n = -5.21$

$\alpha = 0.05 \quad \Rightarrow$ reject when $|Z_n| > z_{\frac{\alpha}{2}} = \underbrace{z_{0.025}}_{1.96}$

Class activity

```r
a <- c(0, 0, 0, -1, 1)
test_stat <- (t(a) %*% coef(m1))/sqrt(t(a) %*% vcov(m1) %*% a)
test_stat
```

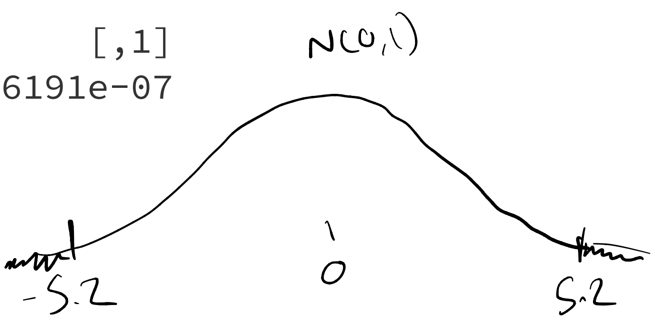
```
##             [,1]
## [1,] -5.207289
```

```r
# rejection region for alpha = 0.05
qnorm(0.025, lower.tail=F)
```

```
## [1] 1.959964
```

```r
# p-value
2*pnorm(abs(test_stat), lower.tail=F)
```

```
##                [,1]
## [1,] 1.916191e-07
```



Logistic regression MLE: $\text{Var}(\hat{\beta}) = V_n = I^{-1}(\beta)$

Summary of Wald tests

Let $\theta \in \mathbb{R}^p$ be some parameter of interest. We wish to test the hypotheses

$$H_0 : C\theta = \gamma_0 \qquad H_A : C\theta \neq \gamma_0$$ Specify hypotheses

for some $C \in \mathbb{R}^{q \times p}$. Given an estimator $\hat{\theta}_n$ such that

$$V_n^{-\frac{1}{2}}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I),$$

the **Wald test** rejects when

$$\underbrace{(C\hat{\theta}_n - \gamma_0)^T (CV_nC^T)^{-1}(C\hat{\theta}_n - \gamma_0)}_{\text{calculate a test statistic}} > \underbrace{\chi^2_{q,\alpha}}_{\text{determine when to reject } H_0}$$

calculate a test statistic

determine when to reject H_0

General framework for hypothesis tests

Definition: Let $\theta \in \mathbb{R}^p$ be some parameter of interest. A **hypothesis** is a statement about θ. The goal of a hypothesis test is to compare two competing hypotheses:

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

- here $\Theta_0 \cap \Theta_1 = \emptyset$
- If $\Theta_0 = \{\theta_0\}$, H_0 is a simple hypothesis

otherwise, H_0 is a composite hypothesis (likewise for H_A)

Example: $H_0 : \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\qquad H_A : \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} \neq 0$

simple null

composite alternative

Outcomes

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

The outcome of the test is a decision to either **reject** H_0 or **fail to reject** H_0.

Four possibilities

		Truth: H_0 is true	Truth: H_A is true
Decision	Fail to reject	Yay!	Type II error (false negative)
	Reject	Type I error (false positive)	Yay!

Goal: 1) Control type I error

2) Minimize type II error

Constructing a test

$$H_0 : \theta \in \Theta_0 \qquad H_A : \theta \in \Theta_1$$

Observe data $X_1, \ldots, X_n$

① Calculate a test statistic $T_n = T(X_1, \ldots, X_n)$

② Choose a rejection region $R = \{(x_1, \ldots, x_n) : \text{reject } H_0\}$

③ Reject H_0 if $(X_1, \ldots, X_n) \in R$

<u>Example</u>: $X_1, \ldots, X_n$ iid from population w/ mean μ, variance σ^2. $H_0: \mu = \mu_0 \quad H_A: \mu \neq \mu_0$

① $T(X_1, \ldots, X_n) = \dfrac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} = Z_n$

② $R = \left\{ (x_1, \ldots, x_n) : \left| \dfrac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \right\}$

③ Reject H_0 if $(X_1, \ldots, X_n) \in R$, i.e. if $|Z_n| > z_{\frac{\alpha}{2}}$

Prob. of a type I error = $P(\text{reject } H_0 \mid H_0 \text{ is true})$

Under H_0, $Z_n \approx N(0,1)$ => Prob. of type I error = $P(|Z| > z_{\frac{\alpha}{2}})$
$= \alpha$

Power function

Suppose we reject H_0 when $(X_1, \ldots, X_n) \in R$. The **power function** $\beta(\theta)$ is

$$\beta(\theta) = P_\theta((X_1, \ldots, X_n) \in R)$$

Example

$X_1, \ldots, X_n$ iid from a population with mean μ and variance σ^2.

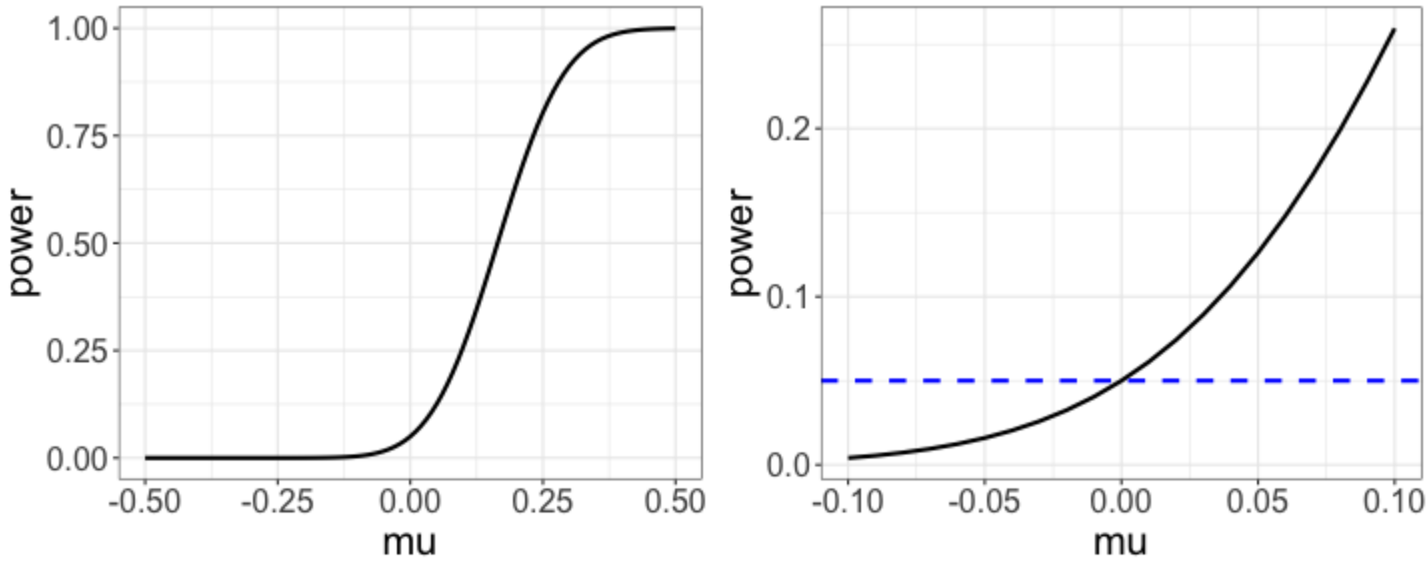
$$H_0 : \mu = \mu_0 \qquad H_A : \mu > \mu_0$$

Class activity

$$\beta(\mu) \approx 1 - \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right)$$

- Suppose that $\mu_0 = 0, n = 100$, and $\sigma = 1$. Make a plot of $\beta(\mu)$ vs. μ for $\alpha = 0.05$.
- Now consider testing $H_0 : \mu \leq \mu_0$ vs. $H_A : \mu > \mu_0$. Will this change our rejection region if we want a size α test?

Class activity



Simulation verification

```
n <- 100
sigma <- 1
mu0 <- 0
nreps <- 5000
test_stats <- rep(0, nreps)
for(i in 1:nreps){
  x <- rnorm(n, mu0, sigma)
  test_stats[i] <- (mean(x) - mu0)/(sigma/sqrt(n))
}
mean(test_stats > qnorm(0.05, lower.tail=F))
```

```
## [1] 0.0548
```

Simulation verification

```
n <- 100
sigma <- 1
mu0 <- 0
nreps <- 5000
test_stats <- rep(0, nreps)
for(i in 1:nreps){
  x <- rnorm(n, 0.1, sigma)
  test_stats[i] <- (mean(x) - mu0)/(sigma/sqrt(n))
}
mean(test_stats > qnorm(0.05, lower.tail=F))
```

```
## [1] 0.2646
```