STAC51 TUT02

Week 3: Jan 28, 2021

Binomial Theorem

Theorem 3.2.2 (Binomial Theorem) *For any real numbers x and y and integer $n \geq 0$,*

(3.2.4) $$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

$\leftarrow$ $\sum_{i=0}^{n} \binom{n}{i} 1^i 1^{(n-i)}$

$x=1$, $y=1$

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n = (1+1)^n$$

Mgf of X ~ binomial(n,p)

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \underbrace{\binom{n}{x} (pe^t)^x (1-p)^{n-x}}$$ use biomonial Thm

$$= [pe^t + (1-p)]^n$$

Multinomial Theorem

$\underline{X} \sim$ Multinomial

$\mathbb{E}(e^{\underline{t}^T \underline{X}})$

$t_1 X_1 + \cdots t_n X_n$

$(X_1, \cdots, X_n)$

Theorem 4.6.4 (Multinomial Theorem) *Let m and n be positive integers. Let $\mathcal{A}$ be the set of vectors $\mathbf{x} = (x_1, \ldots, x_n)$ such that each x_i is a nonnegative integer and $\sum_{i=1}^n x_i = m$. Then, for any real numbers $p_1, \ldots, p_n$,*

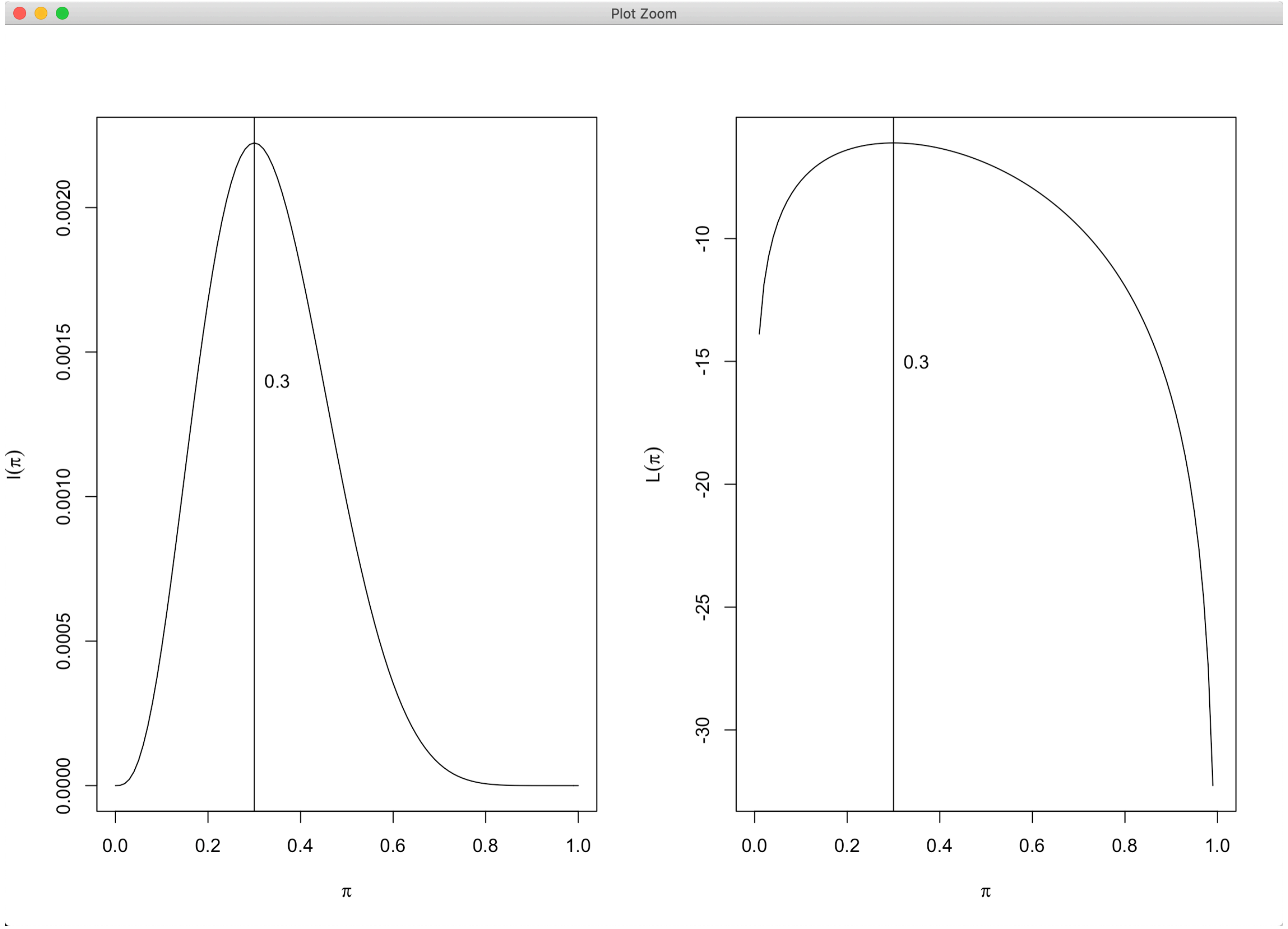
$$(p_1 + \cdots + p_n)^m = \sum_{\mathbf{x} \in \mathcal{A}} \frac{m!}{x_1! \cdot \cdots \cdot x_n!} p_1^{x_1} \cdot \cdots \cdot p_n^{x_n}.$$

$$= \binom{m}{x_1, \cdots, x_n}$$

Binomial Example

- Let's assume that we want to test that a coin is fair
- We toss the coin 10 (n) times and we got 3 (Y) heads
- $\ell(\pi) = \pi^3(1-\pi)^7$ that is $L(\pi) = 3\log(\pi) + 7\log(1-\pi)$. Thus $\hat{\pi} = 3/10$
- Plot the likelihood function and find the value of π that maximizes $\ell(\pi)$.

R Tutorial: The likelihood plots



Example

A coin was tossed 32 times and observed 23 heads. Use the likelihood ratio test to test $H_0 : \pi = 0.5$ and $H_a : \pi \neq 0.5$

$X \sim \text{Binomal}(n=32, \pi)$

$$LRT = -2(\log L(\pi_0) - \log(\hat{\pi})) \sim \chi^2_{(1)}$$

$$T = -2\log(\Lambda) = 2 \sum \text{observed} \left[\log\left(\frac{\text{observed}}{\text{expected}}\right)\right]$$

$$= 2\left[23 \log\left(\frac{23}{32 \times 0.5}\right) + (32-23)\log\left(\frac{32-23}{32 \times 0.5}\right)\right]$$

$$= 6.337 > 3.84 = (1.96)^2 = \chi^2_{(df=1,\ \alpha=0.05)}$$

Reject H_0

$$\Lambda = \frac{L(\pi_0)}{L(\hat{\pi})} = \frac{\pi_0^x (1-\pi_0)^{n-x}}{\hat{\pi}^x (1-\hat{\pi})^{n-x}} = \left(\frac{\pi_0}{\hat{\pi}}\right)^x \left(\frac{1-\pi_0}{1-\hat{\pi}}\right)^{n-x}$$

$$-\log(\Lambda) = x \log\left(\frac{n\hat{\pi}}{n\pi_0}\right) + (n-x)\log\left(\frac{(1-\hat{\pi})n}{(1-\pi)n}\right)$$

$x = 23$

$n = 32$.

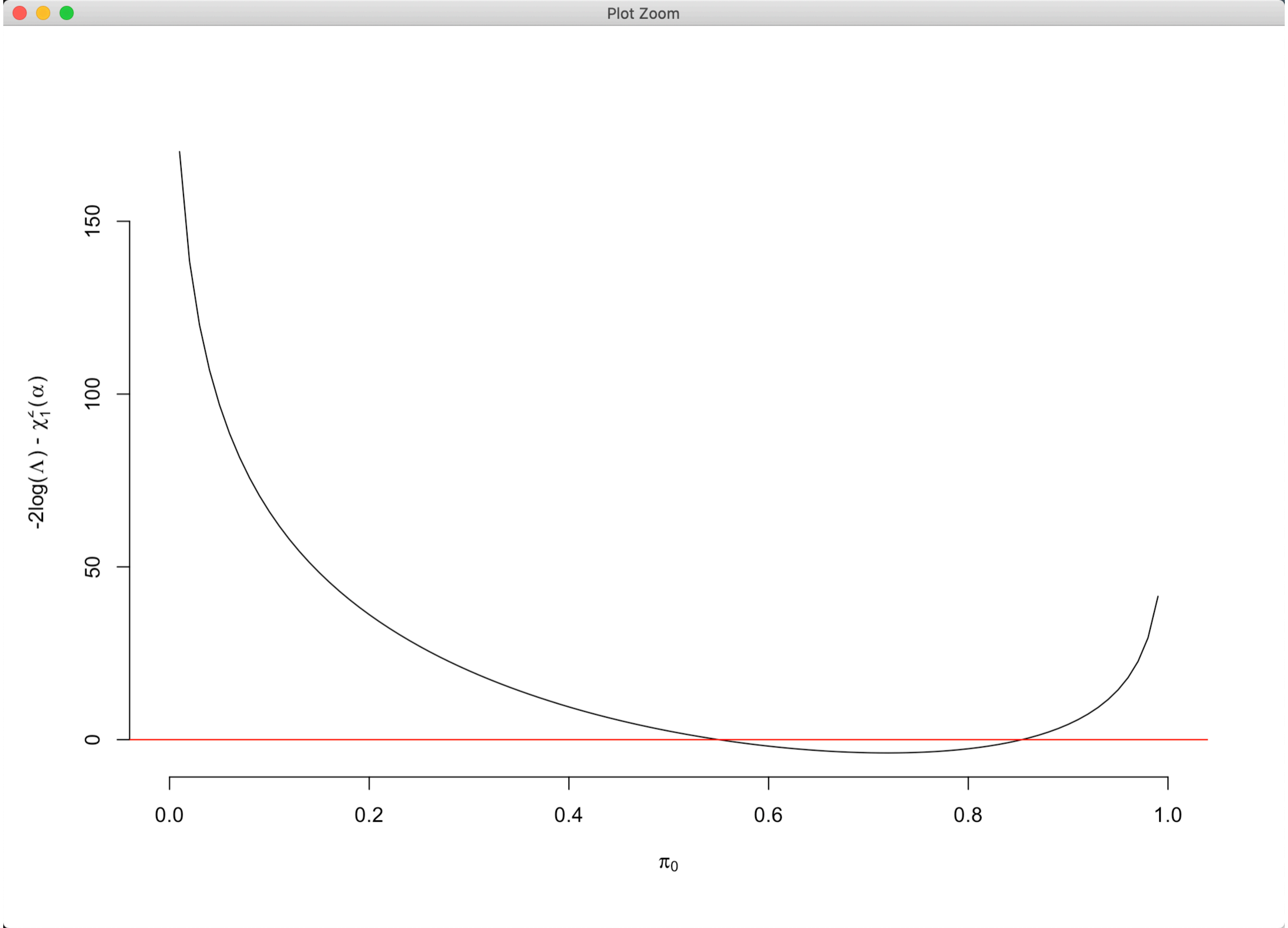
$\hat{\pi} = \frac{23}{32}$.

R Tutorial: LR based Confidence Interval

- Likelihood based confidence interval for π is the set of values of π_0 for which, $-2\log(\Lambda) < \chi_1^2(\alpha)$
- We can find the boundaries of the interval by solving the equation $-2(L_0 - L_1) = \chi_1^2(\alpha)$
- The solution may not be obtainable analytically but we can use numerical methods
- For the previous example we need to solve the following equation,

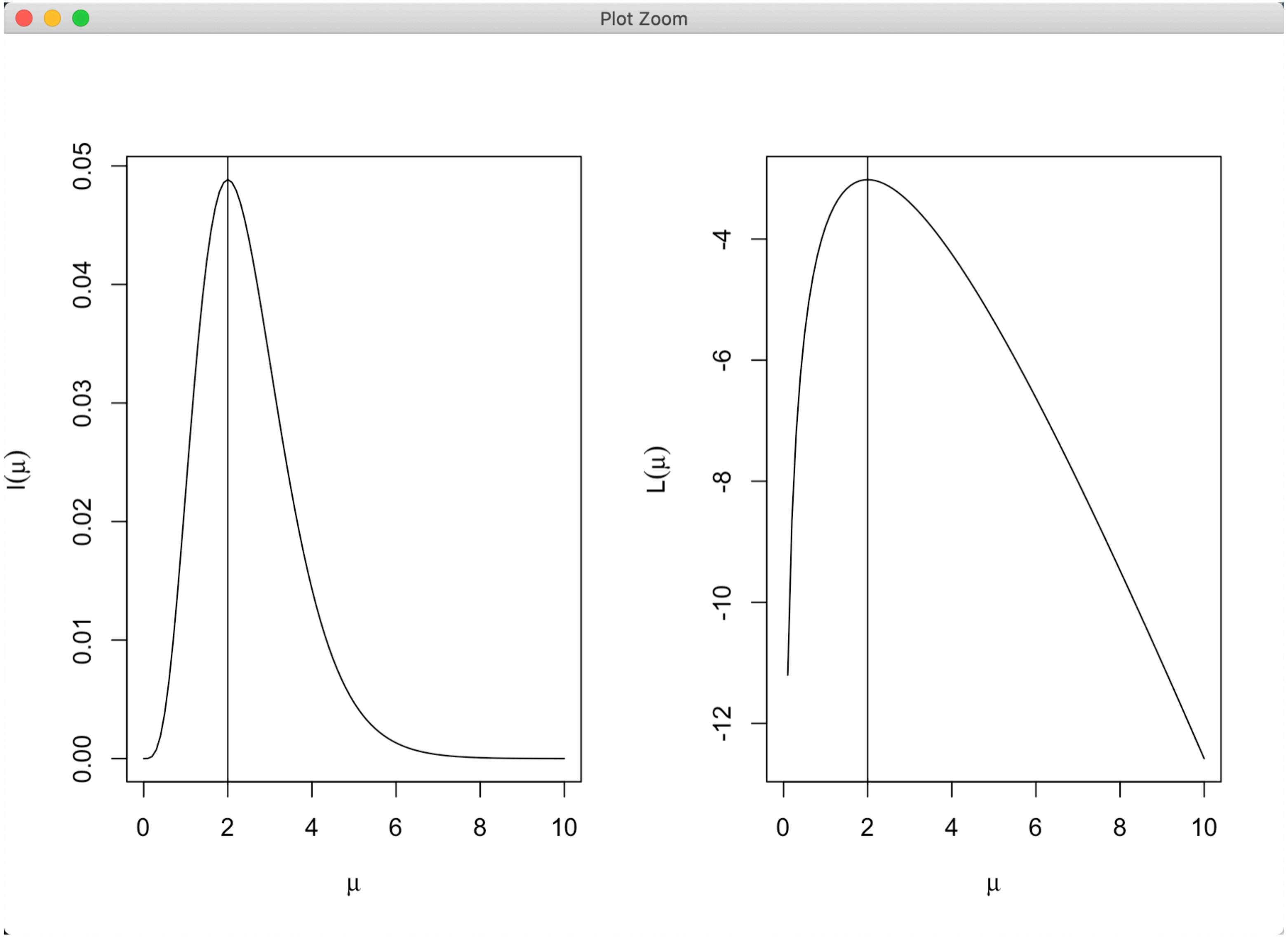
$$2\left[23\log\left(\frac{23}{32\pi_0}\right) + (32-23)\log\left(\frac{32-23}{32-32\pi_0}\right)\right] - 3.84 = 0$$

- We will solve it with `‘R’` in the following slide


Plot Zoom
150
100
50
0
-2log(Λ) - χ²₁(α)
0.0
0.2
0.4
0.6
0.8
1.0
π₀

Assume that the number of cases of tetanus reported in the United States during a single month in 2005 has a Poisson distribution with parameter μ. The number of cases reported in January and February are 1 and 3 respectively.
(a) Find and plot the likelihood function over the space of potential values for μ.
(b) What is the maximum likelihood estimate (MLE) of μ?
(c) Give an estimate of the probability that there is no case of tetanus reported for a given month.

R: plot Poisson likelihood



For $X \sim bin(n,p)$ with $0 < p < 1$, $n \in \mathbb{N}$ we know that

$$\frac{X - np}{\sqrt{npq}} \xrightarrow{d} Z \sim N(0,1) \qquad \text{as} \qquad n \to \infty$$

a) Use the above approximation in distribution plus the fact that for $Z \sim N(0,1)$ we get $P(|Z| \leq 2) \approx .95$ to derive an approximate 95% p-symmetric confidence interval for p and note how its width, W_n, converges to 0 as $n \to \infty$.

$\Leftarrow$ Agaresti Test

b) As $n \to \infty$ we also know that $\widehat{p} = X/n \xrightarrow{d} p$, in which case, by slutsky's theorem, we may amend the denominator of the above approximation, thus to obtain, with much greater ease than in a), a slutsky-amended approximate 95% p-symmetric confidence interval for p. Verify that as $n \to \infty$, the width, V_n, of the amended interval also converges to 0

$\Leftarrow$ WALD Test

- Confidence Interval

$X \sim$ Binomial (n, θ)

$$C(X) = [L(X), U(X)] \Leftarrow \text{Random set.}$$

confidence level

$$\inf_{\theta \in \Theta} P_\theta(\theta \in C(X))$$

$\Downarrow$ Rejection Region

$$\hookrightarrow \beta(\theta) = P_\theta(\theta \notin C(X))$$

$H_0: \pi = \pi_0$ vs. $H_1: \pi > \pi_0$ $\quad | \boxed{\alpha} \; (\pi - \pi_0) |$

① Score Test

$$T = \frac{\sqrt{n}(\hat{\pi} - \pi_0)}{\sqrt{\pi_0(1-\pi_0)}} \longrightarrow N(0,1).$$

Z_α. $\quad \alpha = 0.05. \quad Z_\alpha = 1.68.$

② Reject

$$\frac{\sqrt{n}(\hat{\pi} - \pi_0)}{\sqrt{\pi_0(1-\pi_0)}} > \boxed{Z_\alpha}$$

③ Power

$$P_\pi\left(\hat{\pi} : \frac{\sqrt{n}(\hat{\pi} - \pi_0)}{\sqrt{\pi_0(1-\pi_0)}} > Z_\alpha\right).$$

N $\quad$ P $\quad$ pivotal Distribution

$$\frac{\sqrt{n}[(\hat{\pi} - \pi) + (\pi - \pi_0)]}{\sqrt{\pi(1-\pi)}} \times \frac{\sqrt{\pi(1-\pi)}}{}$$

$\Downarrow$ Z.

STAC58.

Decision

Truth	Reject	Not Reject
H_0	(T1E) α	
H_1	$1-\beta$ φ power	T2E φ β

z-test

$= \pm$

$$P_{H_1}(X \in R) = P_{H_1}(T > Z_\alpha).$$

$\{f_\theta : \theta \in \Theta\}$ Statistical Model

$H_0: \theta = \theta_0$ power $= \beta(\theta) = P_{\theta = \theta_0}(R) = \alpha$.

Size α test $\boxed{\beta(\theta_0) = \alpha.}$ Under H_0.

$P(Z > Z_\alpha)$ $= \beta(\theta)$ Under Ha

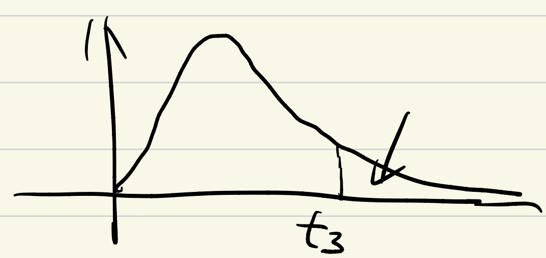
$R(\theta) : \{X :$

$\Rightarrow$ WALD $\dfrac{\sqrt{n}(\hat{\pi} - \pi_0)}{\sqrt{\hat{\pi}(1-\hat{\pi})}} = t_1$ 2x pnorm (obs(t_1), lower tail)

SCORE : $\dfrac{\sqrt{n}(\hat{\pi} - \pi_0)}{\sqrt{\pi_0(1-\pi_0)}} = t_2$

$\chi^2_{(1)}$

LRT : t_3



Example 8.3.2 (Binomial power function) Let $X \sim \text{binomial}(5, \theta)$. Consider testing $H_0 : \theta \le \frac{1}{2}$ versus $H_1 : \theta > \frac{1}{2}$. Consider first the test that rejects H_0 if and only if all "successes" are observed. The power function for this test is

$$\beta_1(\theta) = P_\theta(X \in R) = P_\theta(X = 5) = \theta^5.$$

The graph of $\beta_1(\theta)$ is in Figure 8.3.1. In examining this power function, we might decide that although the probability of a Type I Error is acceptably low ($\beta_1(\theta) \le (\frac{1}{2})^5 = .0312$) for all $\theta \le \frac{1}{2}$, the probability of a Type II Error is too high ($\beta_1(\theta)$ is too small) for most $\theta > \frac{1}{2}$. The probability of a Type II Error is less than $\frac{1}{2}$ only if $\theta > (\frac{1}{2})^{1/5} = .87$. To achieve smaller Type II Error probabilities, we might consider using the test that rejects H_0 if $X = 3$, 4, or 5. The power function for this test is

$$\beta_2(\theta) = P_\theta(X = 3, 4, \text{ or } 5) = \binom{5}{3}\theta^3(1-\theta)^2 + \binom{5}{4}\theta^4(1-\theta)^1 + \binom{5}{5}\theta^5(1-\theta)^0.$$

The graph of $\beta_2(\theta)$ is also in Figure 8.3.1. It can be seen in Figure 8.3.1 that the second test has achieved a smaller Type II Error probability in that $\beta_2(\theta)$ is larger for $\theta > \frac{1}{2}$. But the Type I Error probability is larger for the second test; $\beta_2(\theta)$ is larger for $\theta \le \frac{1}{2}$. If a choice is to be made between these two tests, the researcher must decide which error structure, that described by $\beta_1(\theta)$ or that described by $\beta_2(\theta)$, is more acceptable. ||

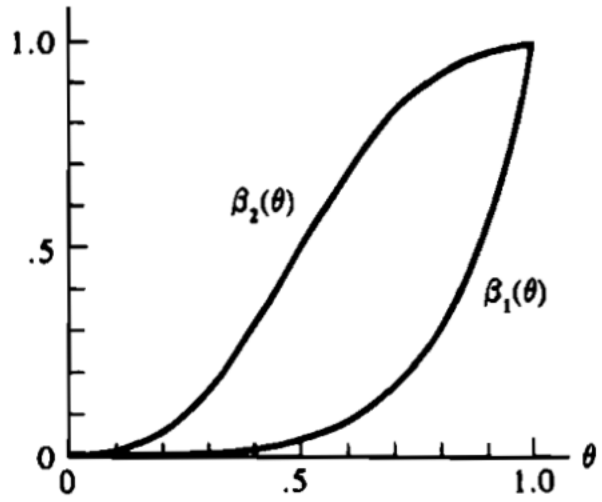


Figure 8.3.1. *Power functions for Example 8.3.2*