LRT

ALRT ⟩ Two recipes for building tests of hypotheses.

only for two-sided hypotheses

STAT 713 sp 2023 Lec 13 slides

$\boxed{\text{Score test}}$ and Wald tests

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Theorem (Asymptotic distribution of Score test statistic)

Let $\mathbf{X}$ *have likelihood* $\mathcal{L}(\theta;\mathbf{X})$ *with score* $S(\theta;\mathbf{X})$ *and Fisher inf.* $I_n(\theta)$ *for* $\theta \in \Theta$, *and consider testing* $H_0\colon \tau(\theta) = \tau_0$ *versus* $H_1\colon \tau(\theta) \neq \tau_0$.

If $\dim(\Theta) = d$ *and* $\Theta_0 = \{\theta \in \Theta : \tau(\theta) = \tau_0\}$ *has dimension* d_0, *then under* H_0

$$S(\hat{\theta}_0;\mathbf{X})^T I_n(\hat{\theta}_0)^{-1} S(\hat{\theta}_0;\mathbf{X}) \xrightarrow{D} \chi^2_{d-d_0}$$

as $n \to \infty$, *where* $\boxed{\hat{\theta}_0 = \operatorname{argmax}_{\theta\in\Theta_0} \mathcal{L}(\theta;\mathbf{X})}$, *under the MLE regularity conditions.*

In consequence, the test with decision rule

$$S(\hat{\theta}_0;\mathbf{X})^T I_n(\hat{\theta}_0)^{-1} S(\hat{\theta}_0;\mathbf{X}) > \chi^2_{d-d_0,\alpha}$$

🙌

has size converging to α *as* $n \to \infty$.

Exercise: Go through heuristics of proof under a simple null $\boxed{H_0\colon \theta = \theta_0.}$

$\hat{\theta}_0 = \theta_0 \qquad \Theta_0 = \{\theta_0\} \qquad d_0 = 0, \quad d = \dim(\Theta).$

Want to show that if $\theta = \theta_0$ (H_0 true) then

$$S(\theta_0; \underset{\sim}{X})^T I_n(\theta_0)^{-1} S(\theta_0; \underset{\sim}{X}) \xrightarrow{D} \chi^2_d \quad \text{as } n \to \infty$$

$$\frac{1}{\sqrt{n}} S(\theta_0; \underset{\sim}{X}) = \frac{1}{\sqrt{n}} \frac{\partial}{\partial\theta} \ell(\theta; \underset{\sim}{X}) \Big|_{\theta=\theta_0}$$

$$= \frac{1}{\sqrt{n}} \frac{\partial}{\partial\theta} \sum_{i=1}^n \log f(x_i; \theta) \Big|_{\theta=\theta_0}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial}{\partial\theta} \log f(x_i; \theta) \Big|_{\theta=\theta_0}$$

$$= \sqrt{n} \, \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\partial}{\partial\theta} \log f(x_i; \theta) \Big|_{\theta=\theta_0} - \mathbb{E} \frac{\partial}{\partial\theta} \log f(x_i; \theta) \Big|_{\theta=\theta_0} \right\}$$

$$\xrightarrow{D} N\left(0, I_1(\theta_0)\right)$$

$$\frac{1}{\sqrt{n}} S(\theta_0; \underset{\sim}{X}) \xrightarrow{D} N\left(0, I_1(\theta_0)\right)$$

$$\Rightarrow \underbrace{\boxed{\frac{1}{\sqrt{n}} \left[I_1(\theta_0)\right]^{-1/2} S(\theta_0; \underset{\sim}{X})}}_{d \times 1} \xrightarrow{D} N(0, I_d)$$

$d \times d$ identity matrix.

$$\left[\frac{1}{\sqrt{n}}\left[I_1(\theta_0)\right]^{-1/2} s(\theta_0;\underset{\sim}{X})\right]^T \frac{1}{\sqrt{n}}\left[I_1(\theta_0)\right]^{-1/2} s(\theta_0;\underset{\sim}{X}) \xrightarrow{D} \chi^2_d$$

$$\left[s(\theta_0;\underset{\sim}{X})\right]^T \frac{1}{\sqrt{n}} I_1(\theta_0)^{-1/2} I_1(\theta_0)^{-1/2} \frac{1}{\sqrt{n}} s(\theta_0;\underset{\sim}{X}) \xrightarrow{D} \chi^2_d$$

$$\left[s(\theta_0;\underset{\sim}{X})\right]^T \frac{1}{n}\left[I_1(\theta_0)\right]^{-1} s(\theta_0;\underset{\sim}{X}) \xrightarrow{D} \chi^2_d$$

$$\left[s(\theta_0;\underset{\sim}{X})\right]^T \left[\underbrace{n I_1(\theta_0)}\right]^{-1} s(\theta_0;\underset{\sim}{X}) \xrightarrow{D} \chi^2_d$$

$$\left[s(\theta_0;\underset{\sim}{X})\right]^T \left[I_n(\theta_0)\right]^{-1} s(\theta_0;\underset{\sim}{X}) \xrightarrow{D} \chi^2_d$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} f_X(x;\beta) = \beta x^{-(\beta+1)}\mathbf{1}(x > 1)$ and consider testing

$$H_0\colon \beta = \beta_0 \text{ versus } H_1\colon \beta \neq \beta_0.$$

Derive the

1. score test.
2. likelihood ratio test. ← takes lots more work
3. asymptotic likelihood ratio test.

$d = 1$

$d_0 = 0$

1

$$L(\beta; \underset{\sim}{X}) = \beta^n \left(\prod_{i=1}^n X_i\right)^{-(\beta+1)}$$

$$\ell(\beta; \underset{\sim}{X}) = n \log \beta - (\beta+1) \sum_{i=1}^n \log X_i$$

$$S(\beta; \underset{\sim}{X}) = \frac{\partial}{\partial \beta} \ell(\beta; \underset{\sim}{X}) = \frac{n}{\beta} - \sum_{i=1}^n \log X_i$$

$$S(\hat{\theta}_0; \mathbf{X})^T I_n(\hat{\theta}_0)^{-1} S(\hat{\theta}_0; \mathbf{X}) > \chi^2_{d-d_0,\alpha}$$

Reject H_0 if
$$\boxed{\frac{[S(\beta_0; \underset{\sim}{X})]^2}{I_n(\beta_0)} > \chi^2_{1,\alpha}}$$

$$I_n(\beta_0) = -\mathbb{E}\left[\frac{\partial}{\partial \beta} S(\beta; \underset{\sim}{X})\right] = -\mathbb{E}\left[-\frac{n}{\beta^2}\right] = \frac{n}{\beta^2}$$

We have

$$\frac{[S(\beta_0; \underset{\sim}{X})]^2}{I_n(\beta_0)} = \frac{\left[\frac{n}{\beta_0} - \sum_{i=1}^{n} \log X_i\right]^2}{n/\beta_0^2}$$

$$= \frac{\beta_0^2}{n}\left[\frac{n}{\beta_0} - \sum \log X_i\right]^2$$

$$= n\,\frac{\beta_0^2}{n^2}\left[\frac{n}{\beta_0} - \sum \log X_i\right]^2$$

$$= n\left[1 - \frac{\beta_0}{n}\sum_{i=1}^{n} \log X_i\right]^2 .$$

Score test rejects $H_0: \beta = \beta_0$ if

$$n\left[1 - \frac{\beta_0}{n}\sum_{i=1}^{n} \log X_i\right]^2 > \chi^2_{1,\alpha} .$$

$$\hat{\beta}_n = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log X_i} \qquad \overset{\|}{\boxed{n\left[1 - \frac{\beta_0}{\underset{\uparrow\ \text{MLE}}{\hat{\beta}_n}}\right]^2 > \chi^2_{1,\alpha}}}$$

② **LRT:**

$$LR(\underline{X}) = \frac{\sup_{\beta \in \{\beta_0\}} L(\beta; \underline{X})}{\sup_{\beta > 0} L(\beta; \underline{X})} = \frac{L(\beta_0; \underline{X})}{L(\hat{\beta}_n; \underline{X})}$$

$$= \frac{\beta_0^{\,n} \left(\prod_{i=1}^n x_i\right)^{-(\beta_0+1)}}{\hat{\beta}_n^{\,n} \left(\prod_{i=1}^n x_i\right)^{-(\hat{\beta}_n+1)}}$$

$$\prod_{i=1}^n x_i = e^{\Sigma \log x_i} = e^{n \cdot \frac{1}{n} \Sigma \log x_i} = e^{n/\hat{\beta}_n}$$

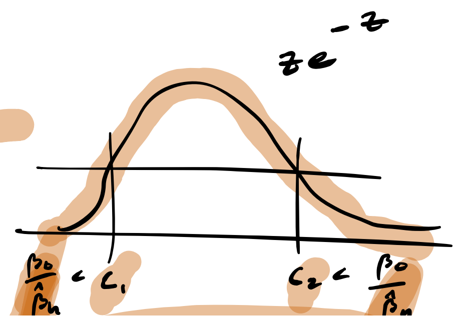
$$= \left(\frac{\beta_0}{\hat{\beta}_n}\right)^n \frac{\left(e^{n/\hat{\beta}_n}\right)^{-(\beta_0+1)}}{\left(e^{n/\hat{\beta}_n}\right)^{-(\hat{\beta}_n+1)}}$$

$$= \left(\frac{\beta_0}{\hat{\beta}_n}\right)^n \frac{\left(e^{-n/\hat{\beta}_n}\right)^{(\beta_0+1)}}{\left(e^{-n/\hat{\beta}_n}\right)^{(\hat{\beta}_n+1)}}$$

$$= \left(\frac{\beta_0}{\hat{\beta}_n}\right)^n \left(e^{-n/\hat{\beta}_n}\right)^{(\beta_0 \cancel{+1}) - (\hat{\beta}_n \cancel{+1})}$$

$$= \boxed{\left(\left(\frac{\beta_0}{\hat{\beta}_n}\right)^n \left(e^{-\frac{\beta_0}{\hat{\beta}_n}}\right)\right)^n} e^n < k$$

reject when



ALRT rejects H_0 when

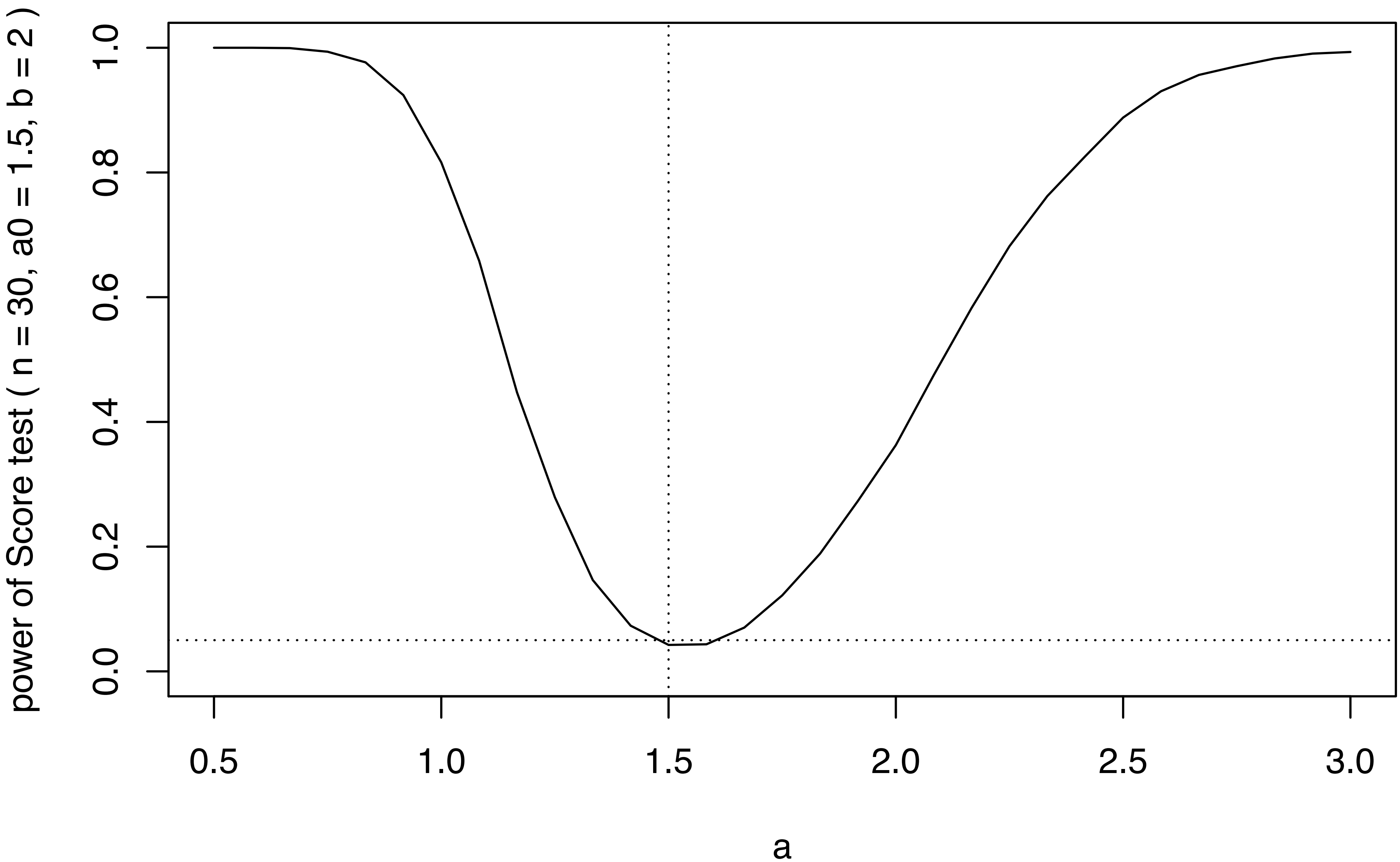
$$-2\log\left[\left(\frac{\beta_0}{\hat{\beta}_n}\right)^n\left(e^{-\frac{\beta_0}{\hat{\beta}_n}}\right)^n e^n\right] > \chi^2_{1,\alpha}$$

$$\Leftrightarrow \quad \frac{\hat{\beta}_n}{\beta_0} > \frac{1}{c_1} \quad \text{or} \quad \frac{\hat{\beta}_n}{\beta_0} < \frac{1}{c_2}$$

Exercise: Consider testing H_0: $a = a_0$ versus H_1: $a \neq a_0$ in the model

$$X_1, \dots, X_n \overset{\text{ind}}{\sim} f_X(x; a, b) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} \exp\left[-\left(\frac{x}{b}\right)^{a}\right] \mathbf{1}(x > 0).$$

1. Derive the score test using $I_1(a, b) = \begin{bmatrix} 1.824/a^2 & -0.423/b \\ -0.423/b & a^2/b^2 \end{bmatrix}$.
2. What is particularly convenient about the score test in this setting?


power of Score test (n = 30, a0 = 1.5, b = 2)
0.0
0.2
0.4
0.6
0.8
1.0
0.5
1.0
1.5
2.0
2.5
3.0
a

Wald test

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0, I_1^{-1}(\theta))$

Delta Method

Let $\hat{\theta}_n$ be the MLE for θ and suppose $\sqrt{n}(\tau(\hat{\theta}_n) - \tau(\theta)) \xrightarrow{D} \text{Normal}(0, \vartheta)$ as $n \to \infty$ for some $\vartheta = \vartheta(\theta)$. Let $\hat{\vartheta}_n$ be any consistent estimator of ϑ and set

$\hat{\vartheta}(\theta) = \vartheta(\hat{\theta}_n)$

$$Z = \sqrt{n}(\tau(\hat{\theta}_n) - \tau_0)/\sqrt{\hat{\vartheta}_n}$$

A Wald-type test has decision rule

1. $|Z| > z_{\alpha/2}$ for testing $H_0\colon \tau(\theta) = \tau_0$ vs $H_1\colon \tau(\theta) \neq \tau_0$.
2. $Z > z_\alpha$ for testing $H_0\colon \tau(\theta) \leq \tau_0$ vs $H_1\colon \tau(\theta) > \tau_0$.
3. $Z < -z_\alpha$ for testing $H_0\colon \tau(\theta) \geq \tau_0$ vs $H_1\colon \tau(\theta) < \tau_0$.

Several choices of $\hat{\vartheta}_n$ are possible:

1. Can use the MLE $\hat{\vartheta}_n = \vartheta(\hat{\theta}_n)$.
2. If τ is 1-1, can use $\hat{\vartheta}_n = \vartheta(\theta_0) = \vartheta(\tau^{-1}(\tau_0))$.
3. Can use bootstrap to estimate ϑ.

Wald tests are not invariant to the function τ (example in hw).

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} f_X(x;\beta) = \beta x^{-(\beta+1)}\mathbf{1}(x > 1)$ and consider testing

$$H_0\colon \beta = \beta_0 \text{ versus } H_1\colon \beta \neq \beta_0.$$

Derive two Wald type tests and run a simulation to get power curves for these and the score, likelihood ratio, and asymptotic likelihood ratio tests.

$$\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{D} N\left(0, \frac{1}{I_1(\beta)}\right), \quad I_n(\beta) = \frac{n}{\beta^2},$$

so $\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{D} N(0, \beta^2)$. $\Rightarrow \dfrac{\sqrt{n}(\hat{\beta}_n - \beta)}{\sqrt{\beta^2}} \xrightarrow{D} N(0,1)$

1. $$Z_1 = \frac{\sqrt{n}(\hat{\beta}_n - \beta_0)}{\sqrt{\hat{\beta}_n^2}} \qquad Z_2 = \frac{\sqrt{n}(\hat{\beta}_n - \beta_0)}{\sqrt{\beta_0^2}}$$

Why plug in β_0?

Under $H_0: \beta = \beta_0$ we have $$\frac{\sqrt{n}\,(\hat{\beta}_n - \beta_0)}{\sqrt{\beta_0^2}} \xrightarrow{D} N(0,1)$$

also $$\frac{\sqrt{n}\,(\hat{\beta}_n - \beta_0)}{\sqrt{\hat{\beta}_n^2}} \xrightarrow{D} N(0,1)$$

Wald 1: Reject H_0 if $|Z_1| > z_{\alpha/2}$

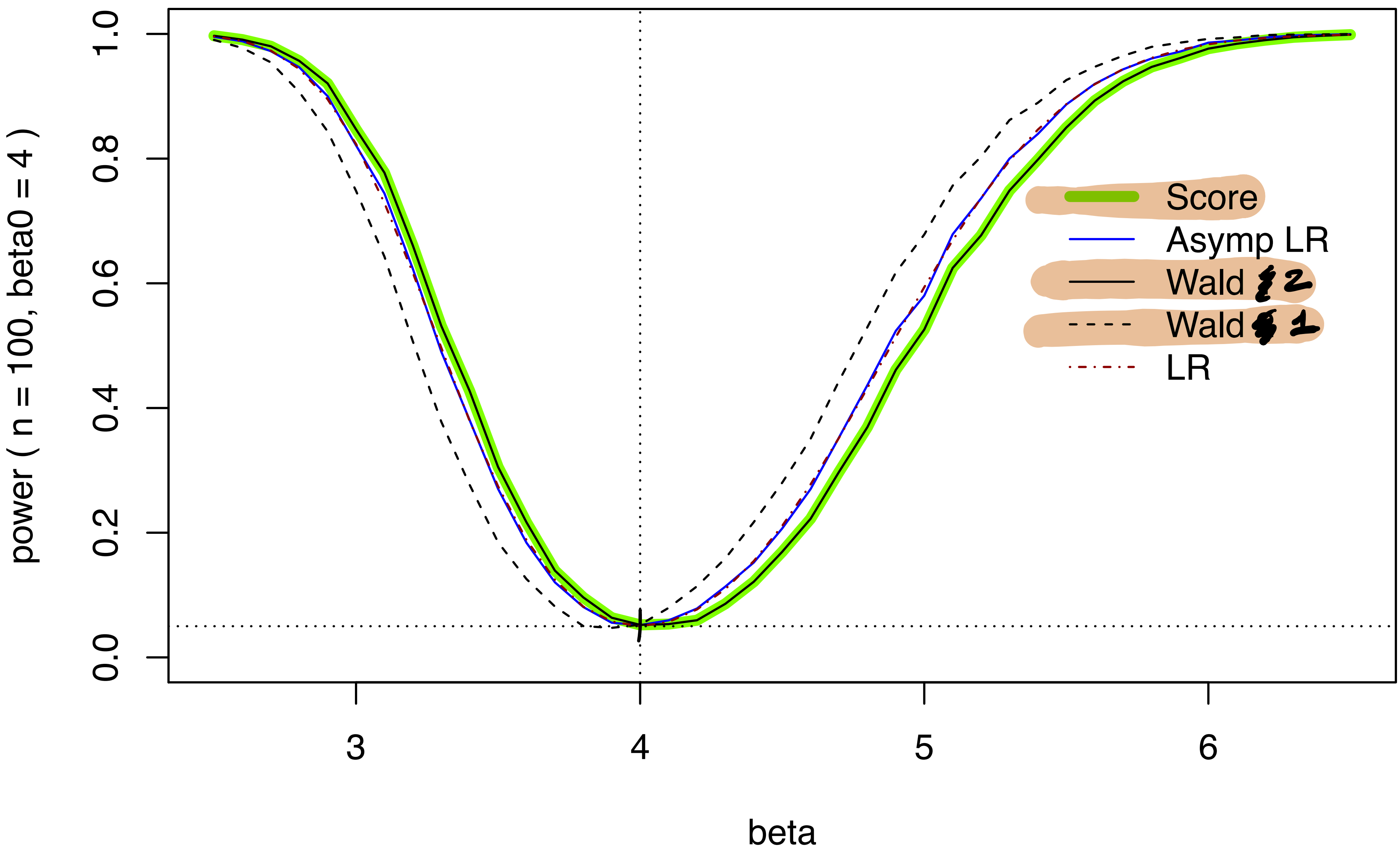
Wald 2: Reject H_0 if $|Z_2| > z_{\alpha/2}$

$$\iff \left| \frac{\sqrt{n}\,(\hat{\beta}_n - \beta_0)}{\sqrt{\beta_0^2}} \right| > z_{\alpha/2}$$

$$\iff \frac{n\,(\hat{\beta}_n - \beta_0)^2}{\beta_0^2} > z_{\alpha/2}^2$$

$$\iff n\left(1 - \frac{\hat{\beta}_n}{\beta_0}\right)^2 > \underbrace{z_{\alpha/2}^2}_{= \chi^2_{1,\alpha}}$$

$\beta_0 = 4$ $\alpha = 0.05$



Under $n = 100$, $\beta_0 = 4$, and $\beta = 4.5$, some realizations are:

	Z1	Z2	ALR	Sc	LR
1	2.80	3.89	9.71	7.84	0.72
2	1.58	1.88	2.81	2.51	0.84
3	1.64	1.96	3.03	2.69	0.84
4	1.70	2.04	3.25	2.87	0.83
5	-0.23	-0.23	0.05	0.05	1.02
6	0.22	0.23	0.05	0.05	0.98
7	1.24	1.42	1.69	1.54	0.88
8	2.24	2.88	5.90	5.00	0.78
9	2.28	2.95	6.14	5.19	0.77
10	-0.07	-0.07	0.00	0.00	1.01
11	2.75	3.79	9.32	7.56	0.73
12	1.35	1.56	2.01	1.83	0.86
13	-0.41	-0.40	0.17	0.17	1.04
14	0.63	0.67	0.41	0.39	0.94
15	-0.40	-0.38	0.16	0.16	1.04

$$Z_1 = \sqrt{n}(\hat{\beta}_n - \beta_0)/\hat{\beta}_n, \quad Z_2 = \sqrt{n}(\hat{\beta}_n - \beta_0)/\beta_0$$

$$\sqrt{n}(\hat{p}_n - p) \xrightarrow{D} N\left(0, \frac{1}{I_1(p)}\right)$$

$$p_X(x) = p^x(1-p)^{1-x} \cdot \mathbb{1}(x \in \{0,1\}) \qquad \Rightarrow \quad \frac{\sqrt{n}(\hat{p}_n - p)}{\sqrt{1/I_1(p)}} \xrightarrow{D} N(0,1)$$

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim}$ Bernoulli(p). For $H_0\colon p = p_0$ vs $H_1\colon p \neq p_0$, consider

replace with $I_1(\hat{p}_n)$

or with $I_1(p_0)$

1. the score test.
2. the likelihood ratio test.
3. the asymptotic likelihood ratio test.
4. the Wald test with Fisher information estimated with the MLE.
5. the Wald test with Fisher information set to its value under H_0.

Compare the power and size of the above tests in a simulation study.

1

$$L(p;\underline{X}) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} = p^{n\bar{X}_n}(1-p)^{n-n\bar{X}_n}$$

$$\ell(p;\underline{X}) = n\bar{X}_n \log p + (n - n\bar{X}_n)\log(1-p)$$

$$S(p;\underline{X}) = \frac{\partial}{\partial p}\ell(p;\underline{X})$$

$$= \frac{n\bar{x}_n}{p} - \frac{(n - n\bar{x}_n)}{1-p}$$

$$= n\bar{x}_n\left[\frac{1}{p} + \frac{1}{1-p}\right] - \frac{n}{1-p}$$

$$= \frac{n\bar{x}_n}{p(1-p)} - \frac{n}{1-p}$$

$$= \frac{n\bar{x}_n - np}{p(1-p)}$$

$$= \frac{n(\bar{x}_n - p)}{p(1-p)}$$

$$I(p;\underset{\sim}{x}) = \mathrm{Var}\left(S(p;\underset{\sim}{x})\right) = n^2\left[\frac{1}{p} + \frac{1}{1-p}\right]^2 \underline{\mathrm{Var}\,\bar{x}_n}$$

$$= n^2\left[\frac{1-\cancel{p} + \cancel{p}}{p(1-p)}\right]^2 \frac{1}{n}\, p(1-p)$$

$$= \frac{n}{p(1-p)}.$$

$H_0: p = p_0$ vs $H_1: p \neq p_0$

Score test statistic

$$S(\underset{\substack{\uparrow\\ p_0}}{\hat{\theta}_0};\mathbf{X})^T I_n(\underset{\substack{\uparrow\\ p_0}}{\hat{\theta}_0})^{-1} S(\underset{\substack{\uparrow\\ p_0}}{\hat{\theta}_0};\mathbf{X}) > \chi^2_{\underset{1-0}{\boxed{d-d_0}},\alpha}$$

$$\frac{[S(p_0;\underset{\sim}{x})]^2}{I_n(p_0)} = \left[\frac{n(\bar{x}_n - p_0)}{p_0(1-p_0)}\right]^2 \Big/ \left[\frac{n}{p_0(1-p_0)}\right] = \frac{n(\bar{x}_n - p_0)^2}{p_0(1-p_0)}$$

Reject H_0 when $\dfrac{n(\bar{x}_n - p_0)^2}{p_0(1-p_0)} > \chi^2_{1,\alpha}$

② $LR(\underset{\sim}{X}) = \dfrac{L(p_0; \underset{\sim}{X})}{L(\hat{p}_n; \underset{\sim}{X})}$

$$= \frac{p_0^{n\bar{x}_n}(1-p_0)^{n-n\bar{x}_n}}{\hat{p}_n^{n\bar{x}_n}(1-\hat{p}_n)^{n-n\bar{x}_n}}$$

$$= \left(\frac{p_0^{\hat{p}_n}}{\hat{p}_n^{\hat{p}_n}}\right)^n \left(\frac{1-p_0}{1-\hat{p}_n}\right)^{n-n\hat{p}_n}$$

Reject H_0 when $\left(\frac{p_0^{\hat{p}_n}}{\hat{p}_n^{\hat{p}_n}}\right)^n \left(\frac{1-p_0}{1-\hat{p}_n}\right)^{n-n\hat{p}_n} < k$

$\iff$ Reject H_0 $\hat{p}_n < c_1$ or $\hat{p}_n > c_2$.

Then could choose c_1, c_2 to get size α.

③ ALRT: Reject H_0

$$-2 \log \left[\left(\frac{p_0^{\hat{p}_n}}{\hat{p}_n^{\hat{p}_n}}\right)^n \left(\frac{1-p_0}{1-\hat{p}_n}\right)^{n-n\hat{p}_n}\right] > \chi^2_{1,\alpha}$$

④ Wald with MLE estimate of standard error:

$Z_1 = \sqrt{n}(\hat{p}_n - p_0) \Big/ \sqrt{1/I_1(\hat{p}_n)}$ $\qquad I_1(p) = \frac{1}{p(1-p)}$

under $H_0: p = p_0$

$$= \frac{\sqrt{n}\,(\hat{p}_n - p_0)}{\sqrt{\hat{p}_n(1-\hat{p}_n)}} \xrightarrow{D} N(0,1)$$

Reject when $|Z_1| > z_{\alpha/2}$

(5) Wald with null value of standard error

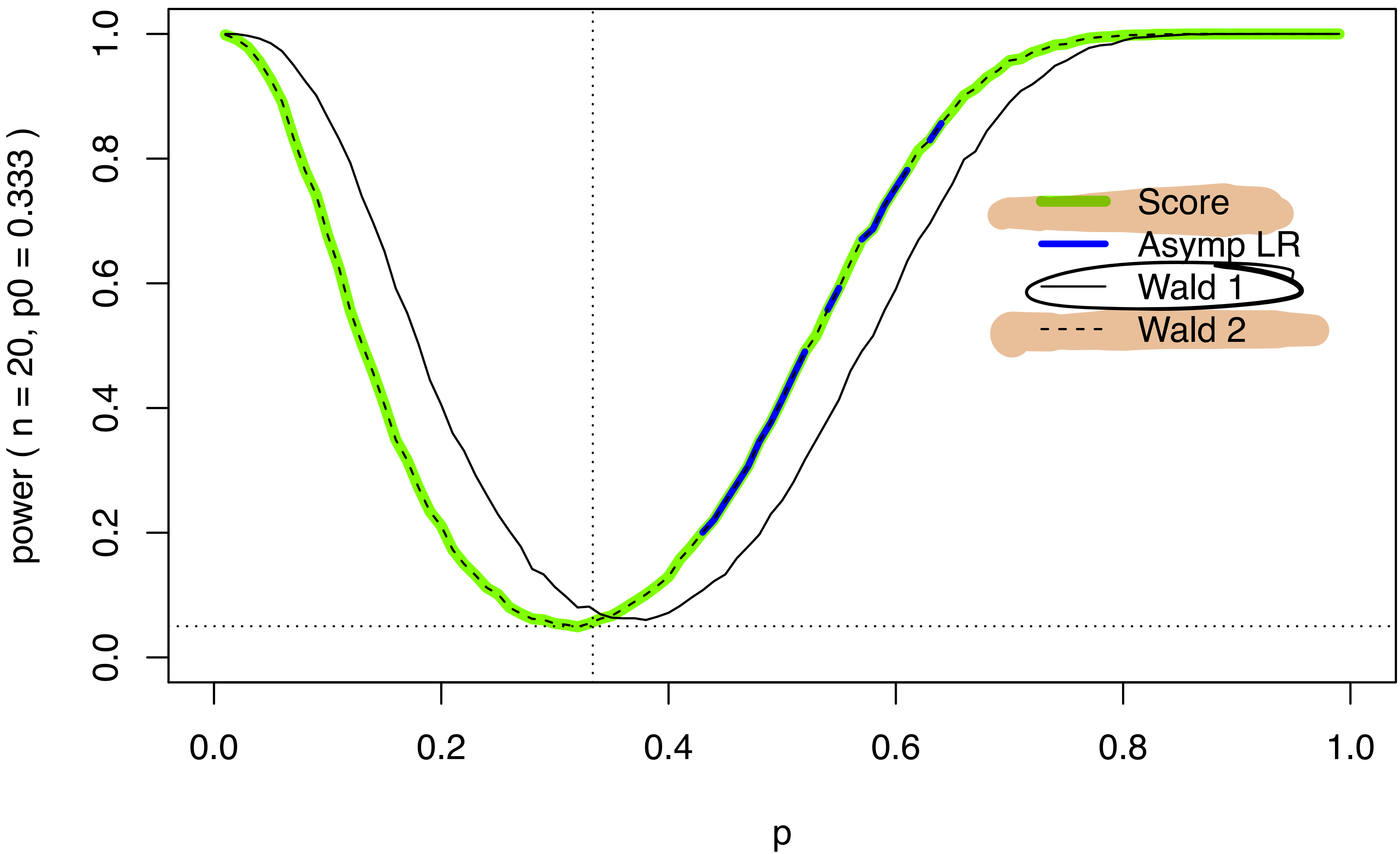
Under $H_0: p = p_0$

$$Z_2 = \frac{\sqrt{n}\,(\hat{p}_n - p_0)}{\sqrt{p_0(1-p_0)}} \xrightarrow{D} N(0,1).$$

Reject when $|Z_2| > z_{\alpha/2}$

Score test:

$$\underbrace{\frac{n\,(\bar{X}_n - p_0)^2}{p_0(1-p_0)}}_{= Z_2^2} > \chi^2_{1,\alpha}$$

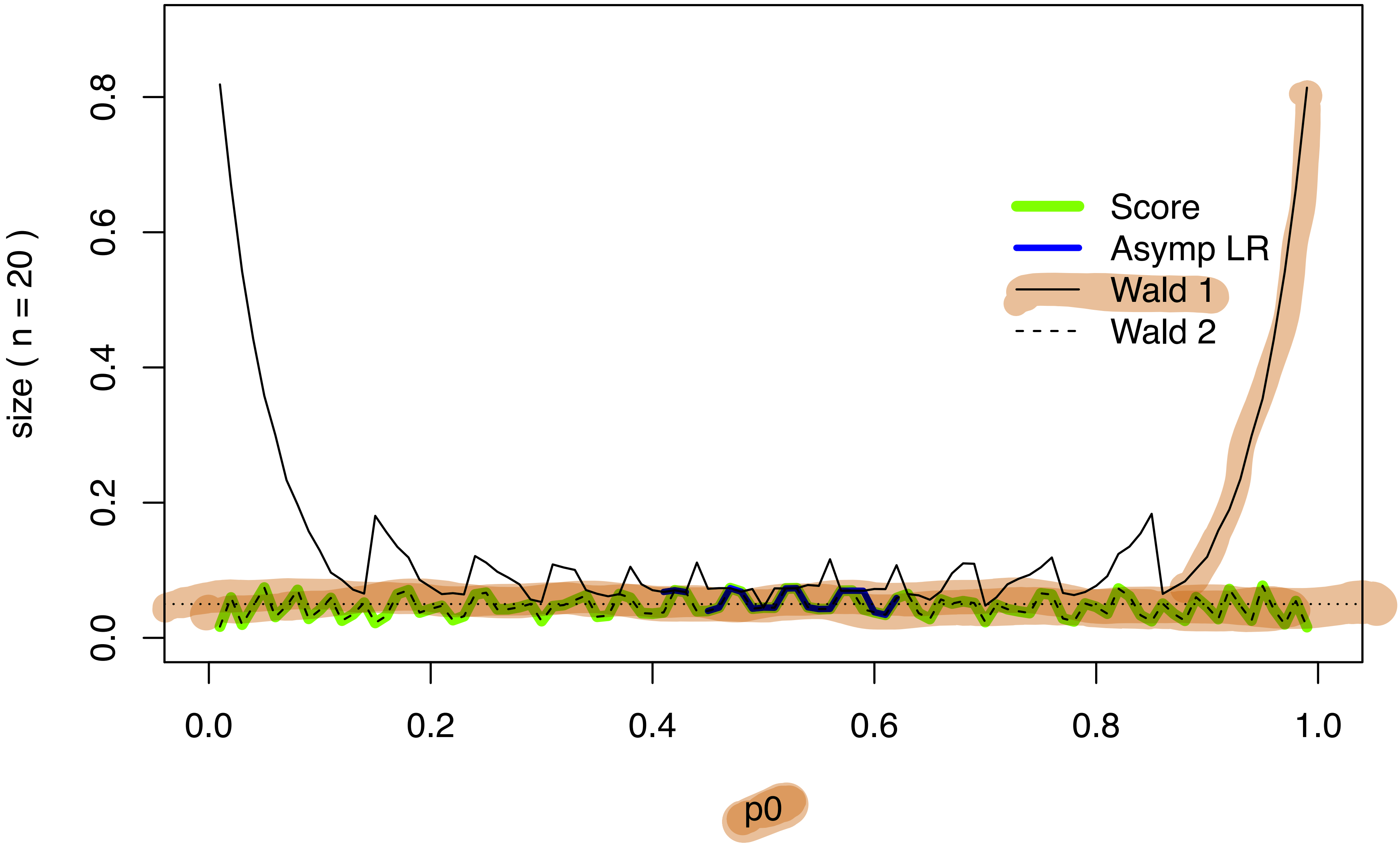


power (n = 20, p0 = 0.333)
1.0
0.8
0.6
0.4
0.2
0.0
Score
Asymp LR
Wald 1
Wald 2
0.0
0.2
0.4
0.6
0.8
1.0
p

Under $n = 20$, $p_0 = 1/3$, and $p = 1/4$, some realizations are:

	Z_1	Z_2	ALR	Sc
1	0.61	0.63	0.39	0.40
2	-0.33	-0.32	0.10	0.10
3	-1.49	-1.26	1.75	1.60
4	-0.86	-0.79	0.66	0.62
5	0.16	0.16	0.02	0.02
6	-1.49	-1.26	1.75	1.60
7	-0.86	-0.79	0.66	0.62
8	-5.81	-2.69	9.66	7.22
9	-0.86	-0.79	0.66	0.62
10	-0.33	-0.32	0.10	0.10
11	-1.49	-1.26	1.75	1.60
12	-0.33	-0.32	0.10	0.10
13	-2.30	-1.74	3.47	3.02
14	-0.86	-0.79	0.66	0.62
15	-0.86	-0.79	0.66	0.62

$$Z_1 = \sqrt{n}(\hat{p}_n - p_0)/\sqrt{\hat{p}_n(1 - \hat{p}_n)}, \quad Z_2 = \sqrt{n}(\hat{p}_n - p_0)/\sqrt{p_0(1 - p_0)}$$


size (n = 20)
0.0
0.2
0.4
0.6
0.8
p0
1.0
Score
Asymp LR
Wald 1
Wald 2