STAT 713 sp 2023 Lec 13 slides

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* $\hat\theta_0 = \operatorname{argmax}_{\theta\in\Theta_0} \mathcal{L}(\theta;\mathbf{X})$, *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 $H_0\colon \theta = \theta_0$.

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

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

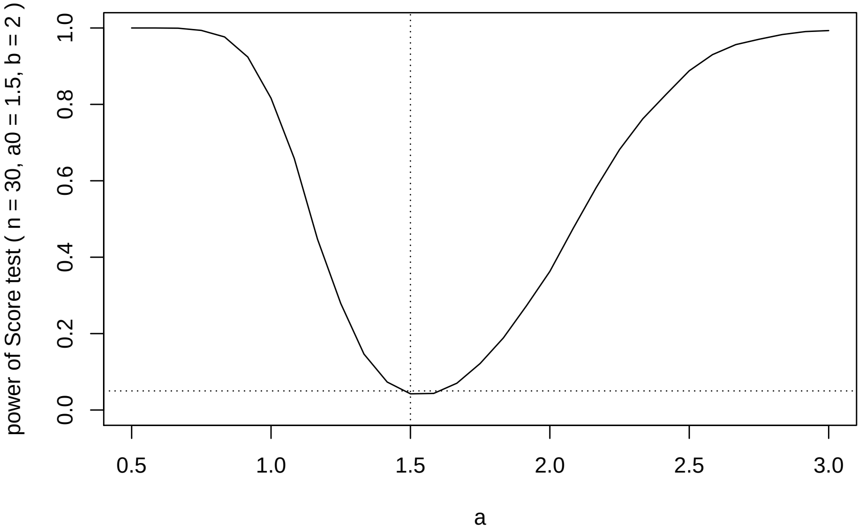
Derive the

1. score test.
2. likelihood ratio test.
3. asymptotic likelihood ratio test.

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

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

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

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:

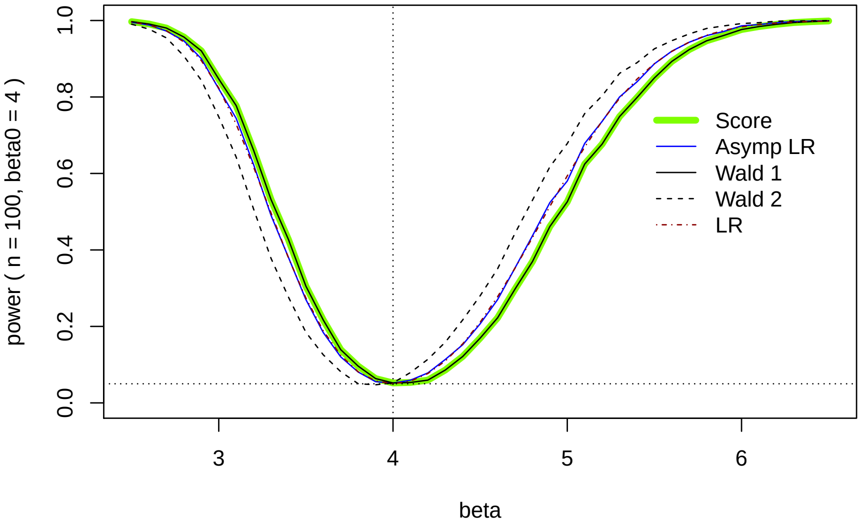
1. Can use the MLE $\hat{\vartheta}_n = \vartheta(\hat{\theta}_n)$.
2. If τ is 1-1, can use $\hat{\vartheta}_n = \vartheta(\theta_0) = \vartheta(\tau^{-1}(\tau_0))$.
3. Can use bootstrap to estimate ϑ.

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.


power (n = 100, beta0 = 4)
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
beta
Score
Asymp LR
Wald 1
Wald 2
LR

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

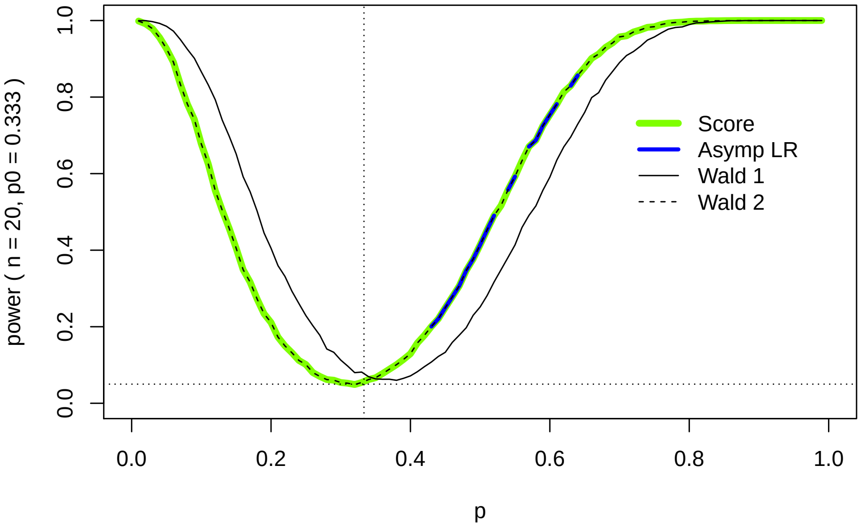
	Z1	Z2	ALR	Sc	LR
1	2.80	3.89	9.71	7.84	0.72
2	1.58	1.88	2.81	2.51	0.84
3	1.64	1.96	3.03	2.69	0.84
4	1.70	2.04	3.25	2.87	0.83
5	-0.23	-0.23	0.05	0.05	1.02
6	0.22	0.23	0.05	0.05	0.98
7	1.24	1.42	1.69	1.54	0.88
8	2.24	2.88	5.90	5.00	0.78
9	2.28	2.95	6.14	5.19	0.77
10	-0.07	-0.07	0.00	0.00	1.01
11	2.75	3.79	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

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

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

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.

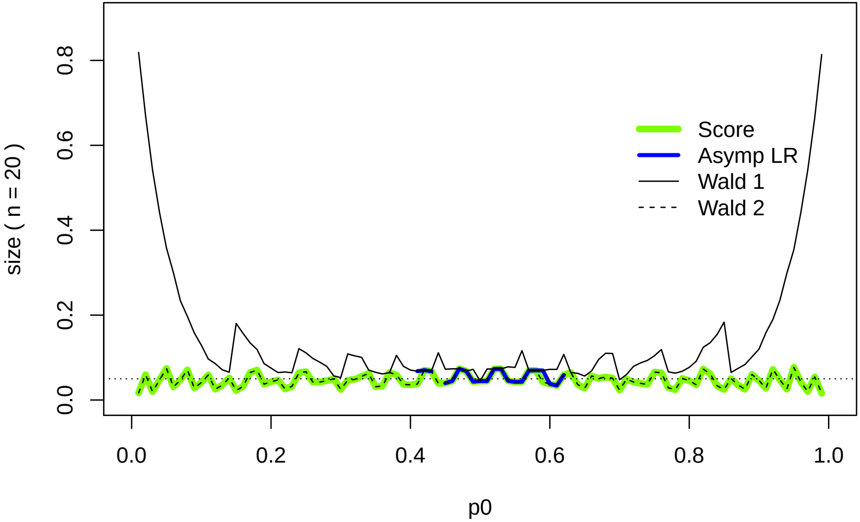
Compare the power and size of the above tests in a simulation study.


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