N-P $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$ found UMP

K-R $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$ found UMP

STAT 713 sp 2023 Lec 12 slides

Likelihood ratio tests → $H_0: \theta \in \Theta_0$

$H_1: \theta \in \Theta_1$

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.

Likelihood ratio

For $\mathbf{X}$ having likelihood $\mathcal{L}(\theta; \mathbf{X})$ and for some hypotheses

$$H_0\colon \theta \in \Theta_0 \text{ versus } H_1\colon \theta \in \Theta_1,$$

the *likelihood ratio (LR)* is defined as

$$\mathrm{LR}(\mathbf{X}) = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; \mathbf{X})}{\sup_{\theta \in \Theta} \mathcal{L}(\theta; \mathbf{X})}.$$

The likelihood ratio must take values in the interval $[0, 1]$.

A smaller (larger/smaller) likelihood ratio casts more (more/less) doubt on H_0.

Likelihood ratio test

A *likelihood ratio test (LRT)* is a test which rejects H_0 if $\mathrm{LR}(\mathbf{X}) < k$ for some k.

A <u>larger</u> value of k gives the test <u>greater</u> power over over the alternate space and <u>greater</u> size.

The critical value k can be chosen to give the test a desired size.

Karlin-Rubin: LRTs are UMP for one-sided tests when there is a one-dimensional parameter and a one-dimensional sufficient statistic.

$$L(\mu;\underset{\sim}{X}) = (2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right]$$

Exercise: Let $X_1,\dots,X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu,\sigma^2)$, σ^2 known. For

$$H_0\colon \underbrace{\boxed{\mu = \mu_0}}_{\mu \in \{\mu_0\}} \text{ versus } H_1\colon \mu \neq \mu_0$$

1. Give the LR.
2. Calibrate the LRT to have size α for any $\alpha \in (0,1)$.

$$LR(\underset{\sim}{X}) = \frac{\sup_{\mu\in\{\mu_0\}} L(\mu,\underset{\sim}{X})}{\sup_{\mu\in\mathbb{R}} L(\mu,\underset{\sim}{X})} = \frac{L(\mu_0;\underset{\sim}{X})}{L(\hat{\mu};\underset{\sim}{X})}$$

$$\text{, where } \hat{\mu} = \underset{\mu\in\mathbb{R}}{\operatorname{argmax}}\, L(\mu;\underset{\sim}{X}) \ (\text{MLE}) = \bar{X}_n$$

$$\frac{L(\mu_0;\underset{\sim}{X})}{L(\hat{\mu};\underset{\sim}{X})} = \frac{\cancel{(2\pi)^{-n/2}}\,\cancel{(\sigma^2)^{-n/2}}\,\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_0)^2\right]}{\cancel{(2\pi)^{-n/2}}\,\cancel{(\sigma^2)^{-n/2}}\,\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\bar{x})^2\right]}$$

$$= \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x}_n)^2\right)\right]$$

Reject $H_0: \mu=\mu_0$ if $LR(\underset{\sim}{X}) < k$

$\Longleftrightarrow$

$$\exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x}_n)^2\right)\right] < k$$

$$\Longleftrightarrow \quad \sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x}_n)^2 > -2\sigma^2\log k$$

$\Longleftrightarrow$

$$\sum_{i=1}^{n}\left((x_i-\bar{x}_n)+(\bar{x}_n-\mu_0)\right)^2 - \sum_{i=1}^{n}(x_i-\bar{x}_n)^2 > -2\sigma^2\log k$$

$$\Longleftrightarrow \quad \sum(x_i-\bar{x}_n)^2 + n(\bar{x}_n-\mu_0)^2 - \sum(x_i-\bar{x}_n)^2 > -2\sigma^2\log k$$

$$\Longleftrightarrow \quad \frac{n(\bar{x}_n-\mu_0)^2}{\sigma^2} > -2\log k$$

$$\Longleftrightarrow \quad \frac{\sqrt{n}\,|\bar{x}_n-\mu_0|}{\sigma} > \underbrace{\sqrt{-2\log k}}_{k^*}$$

(ii)

$$\gamma(\mu) = P_{\mu}\left(\frac{\sqrt{n}\,|\bar{X}_n - \mu_0|}{\sigma} > k^* \right)$$

$$= 1 - P_{\mu}\left(\frac{\sqrt{n}\,|\bar{X}_n - \mu_0|}{\sigma} \leq k^* \right)$$

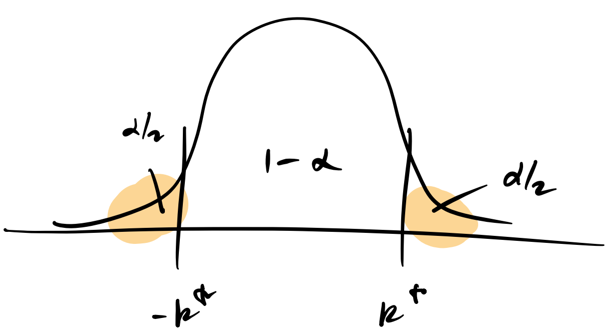
$$= 1 - P_{\mu}\left(-k^* \leq \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma} \leq k^* \right)$$

$$\text{size} = \sup_{\mu \in \{\mu_0\}} \gamma(\mu)$$

$$= \gamma(\mu_0)$$

$$= 1 - P_{\mu_0}\left(-k^* \leq \underbrace{\frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma}}_{Z \sim N(0,1)} \leq k^* \right)$$

$$= 1 - P(-k^* \leq Z \leq k^*) \overset{\text{set}}{=} \alpha$$



$$\Rightarrow \quad k^* = z_{\alpha/2}.$$

$$\Rightarrow \quad \boxed{\text{Reject } H_0: \mu = \mu_0 \text{ if } \frac{\sqrt{n}\,|\bar{X}_n - \mu|}{\sigma} > z_{\alpha/2}.}$$

Finding the likelihood ratio

$H_0: \theta \in \Theta_0$ vs $H_1: \theta \in \Theta_1$

We can write the likelihood ratio as

$$\boxed{\mathrm{LR}(\mathbf{X})} = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; \mathbf{X})}{\sup_{\theta \in \Theta} \mathcal{L}(\theta; \mathbf{X})} = \boxed{\frac{\mathcal{L}(\hat{\theta}_0; \mathbf{X})}{\mathcal{L}(\hat{\theta}; \mathbf{X})}},$$

where

$$\hat{\theta}_0 = \mathrm{argmax}_{\theta \in \Theta_0}\, \ell(\theta; \mathbf{X})$$
$$\hat{\theta} = \mathrm{argmax}_{\theta \in \Theta}\, \ell(\theta; \mathbf{X})$$

So we just need to find these and plug them in:

- $\hat{\theta}_0$ is a *restricted maximum likelihood estimator*; best estimator in null space.
- $\hat{\theta}$ is the MLE.

Note: If the null is a simple hypothesis, i.e. H_0: $\theta = \theta_0$, then $\hat{\theta}_0 = \theta_0$.

Also: If $\hat{\theta} \in \Theta_0$, then $\hat{\theta} = \hat{\theta}_0$, so that $\mathrm{LR}(\mathbf{X}) = 1$.

$$L(\mu; \underset{\sim}{X}) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \right]$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, σ^2 known. Consider testing

$$H_0\colon \mu \leq \mu_0 \text{ versus } H_1\colon \mu > \mu_0.$$

1. Give the LR.
2. Calibrate the LRT to have size α for any $\alpha \in (0, 1)$.

① $$LR(\underset{\sim}{X}) = \frac{\sup_{\mu \leq \mu_0} L(\mu; \underset{\sim}{X})}{\sup_{\mu \in \mathbb{R}} L(\mu; \underset{\sim}{X})} = \frac{L(\hat{\mu}_0; \underset{\sim}{X})}{L(\hat{\mu}; \underset{\sim}{X})}$$

$\bar{X}_n$

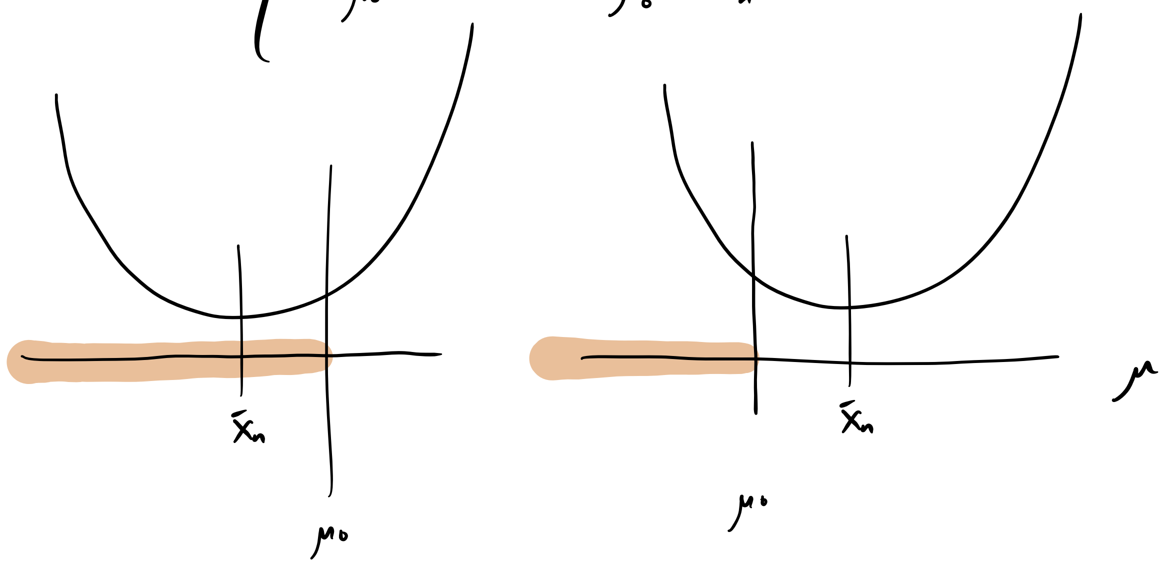
$$\hat{\mu}_0 = \underset{\mu \le \mu_0}{\arg\max}\ L(\mu; \underset{\sim}{x})$$

$$= \underset{\mu \le \mu_0}{\arg\max}\ (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right]$$

$$= \underset{\mu \le \mu_0}{\arg\min}\ \sum_{i=1}^{n} (x_i - \mu)^2 \quad \Leftarrow \text{ a parabola in } \mu$$

$$= \underset{\mu \le \mu_0}{\arg\min} \left\{ \sum_{i=1}^{n} x_i^2 - 2 n \bar{x}_n \mu + n\mu^2 \right\}$$

$$= \begin{cases} \bar{x}_n & \text{if } \mu_0 \ge \bar{x}_n \\ \mu_0 & \text{if } \mu_0 < \bar{x}_n \end{cases}$$



$H_0: \mu \le \mu_0$ vs $H_1: \mu > \mu_0$

$$LR(\underset{\sim}{x}) = \frac{L(\hat{\mu}_0; \underset{\sim}{x})}{L(\hat{\mu}; \underset{\sim}{x})} = \begin{cases} \dfrac{L(\bar{x}_n; \underset{\sim}{x})}{L(\bar{x}_n; \underset{\sim}{x})} & \text{if } \bar{x}_n \le \mu_0 \\[2ex] \dfrac{L(\mu_0; \underset{\sim}{x})}{L(\bar{x}_n; \underset{\sim}{x})} & \text{if } \bar{x}_n > \mu_0 \end{cases}$$

$$= \begin{cases} 1 & \text{if } \bar{x}_n \le \mu_0 \quad \left(\begin{matrix}\hat{\mu} \le \mu_0 \\ \hat{\theta} \in \Theta_0\end{matrix}\right) \\[2ex] \dfrac{L(\mu_0; \underset{\sim}{x})}{L(\bar{x}_n; \underset{\sim}{x})} & \text{if } \bar{x}_n > \mu_0 \quad \left(\begin{matrix}\hat{\mu} > \mu_0 \\ \hat{\theta} \in \Theta_1\end{matrix}\right) \end{cases}$$

$$\frac{L(\mu_0; \underset{\sim}{x})}{L(\bar{x}_n; \underset{\sim}{x})} = \frac{(2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 \right]}{(2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \right]}$$

$$= \frac{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 \right]}{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \right]}.$$

LRT says reject H_0 when $\dfrac{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 \right]}{\exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \right]} < k.$

(provided $\bar{x}_n > \mu_0$)

$\iff$

$$\exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}(x_i-\mu_0)^2-\sum_{i=1}^{n}(x_i-\bar{x}_n)^2\right)\right] < k$$

$\iff$

$$\sum_{i=1}^{n}(x_i-\mu_0)^2-\sum_{i=1}^{n}(x_i-\bar{x}_n)^2 > -2\sigma^2 \log k$$

$\iff$

$\vdots$

$$\iff \quad \frac{\sqrt{n}(\bar{x}_n-\mu_0)}{\sigma} > k^{*}$$

Exercise: Let $Y \sim \text{Geometric}(p)$ distribution, $p \in [0, 1]$ unknown. Consider testing

$$H_0\colon p \leq p_0 \text{ versus } H_1\colon p > p_0.$$

1. Give the LR.
2. Give a level-α likelihood ratio test.

1. LRTs for Normal mean and variance

2. The asymptotic likelihood ratio test

$\theta = (\mu, \sigma^2)$

$\Theta = \mathbb{R} \times (0, \infty)$

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, μ and σ^2 unknown, and consider

$$H_0\colon \mu \leq \mu_0 \text{ versus } H_1\colon \mu > \mu_0.$$

Show that the test

$$\text{Reject } H_0 \text{ iff } \sqrt{n}(\bar{X}_n - \mu_0)/S_n > t_{n-1,\alpha}$$

is the LRT with size α. Steps:

1. Find the MLEs $\hat{\mu}$ and $\hat{\sigma}^2$. $\hat{\mu} = \bar{x}_n$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x}_n)^2$
2. Find the restricted MLEs $\hat{\mu}_0$ and $\hat{\sigma}_0^2$ under $H_0\colon \mu \leq \mu_0$.
3. Show that $\text{LR} < c$ is equivalent to $\sqrt{n}(\bar{X}_n - \mu_0)/S_n > c_1$.

Restricted MLEs

$$LR(\underset{\sim}{X}) = \frac{\sup_{\mu \leq \mu_0, \sigma^2 > 0} L(\mu, \sigma^2; \underset{\sim}{X})}{\sup_{\mu \in \mathbb{R}, \sigma^2 > 0} L(\mu, \sigma^2; \underset{\sim}{X})} = \frac{L(\hat{\mu}_0, \hat{\sigma}_0^2; \underset{\sim}{X})}{L(\hat{\mu}, \hat{\sigma}^2; \underset{\sim}{X})}$$

MLES

$$\ell(\mu, \sigma^2; \underset{\sim}{X}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{\sum_{i=1}^n (X_i-\mu)^2}{2\sigma^2}$$

Find, for any μ, the maximizer in σ^2.

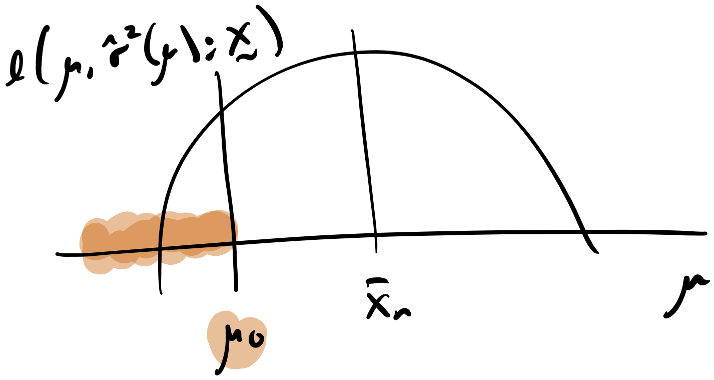
$$\frac{\partial}{\partial\sigma^2}\ell(\mu, \sigma^2; \underset{\sim}{X}) = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n (X_i-\mu)^2}{2(\sigma^2)^2} \overset{set}{=} 0$$

$$\rightsquigarrow \quad \hat{\sigma}^2(\mu) = \frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2$$

$$\ell(\mu, \hat{\sigma}^2(\mu); \underset{\sim}{X}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2\right)$$

$$- \frac{\sum_{i=1}^n (X_i-\mu)^2}{2\,\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2}$$

$$= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\Big(\underbrace{\frac{1}{n}\sum_{i=1}^n (X_i-\mu)^2}_{\text{parabola in } \mu}\Big) - \frac{n}{2}$$



$$\hat{\mu}_0 = \underset{\mu \le \mu_0}{\operatorname{argmax}}\ \ell(\mu, \hat{\sigma}^2(\mu); \underset{\sim}{X}) = \begin{cases} \mu_0 & \bar{X}_n > \mu_0 \\ \bar{X}_n & \bar{X}_n \le \mu_0 \end{cases}$$

$$\hat{\sigma}_0^2 = \hat{\sigma}^2(\hat{\mu}_0) = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2 & \bar{X}_n > \mu_0 \\ \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 & \bar{X}_n \le \mu_0 \end{cases}$$

$$\hat{\sigma}^2(\mu) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

$$(\hat{\mu}_0, \hat{\sigma}_0^2) = \begin{cases} \left(\mu_0, \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2\right) & \underline{\bar{X}_n > \mu_0} \\ \left(\bar{X}_n, \underline{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}\right) & \bar{X}_n \le \mu_0 \end{cases}$$

$LR(\underset{\sim}{x}) =$

$$\frac{L(\hat{\mu}_0, \hat{\sigma}_0^2; \underset{\sim}{x})}{L(\hat{\mu}, \hat{\sigma}^2; \underset{\sim}{x})} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\hat{\sigma}_0}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i - \hat{\mu}_0)^2}{2\sigma_0^2}\right]}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\hat{\sigma}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{2\sigma^2}\right]}$$

$$= \begin{cases} \dfrac{\left(\frac{1}{\sqrt{\frac{1}{n}\sum(x_i-\mu_0)^2}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2\,\frac{1}{n}\sum(x_i-\mu_0)^2}\right]}{\left(\frac{1}{\sqrt{\frac{1}{n}\sum(x_i-\bar{x}_n)^2}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{2\,\frac{1}{n}\sum(x_i-\bar{x}_n)^2}\right]} & \bar{x}_n > \mu_0 \\ 1 & \bar{x}_n \le \mu_0 \end{cases}$$

$$= \begin{cases} \left(\dfrac{\sum_{i=1}^n (x_i-\bar{x}_n)^2}{\sum_{i=1}^n (x_i-\mu_0)^2}\right)^{n/2}, & \bar{x}_n > \mu_0 \\ 1 & \bar{x}_n \le \mu_0 \end{cases}$$

So, if $\bar{x}_n > \mu_0$, reject $H_0: \mu \le \mu_0$ if

$$\left(\frac{\sum_{i=1}^n (x_i-\bar{x}_n)^2}{\sum_{i=1}^n (x_i-\mu_0)^2}\right)^{n/2} < k$$

$\iff$

$$\boxed{\frac{\sqrt{n}\,(\bar{X}_n - \mu_0)}{S_n} > c} \quad \text{for some } c,$$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

Power

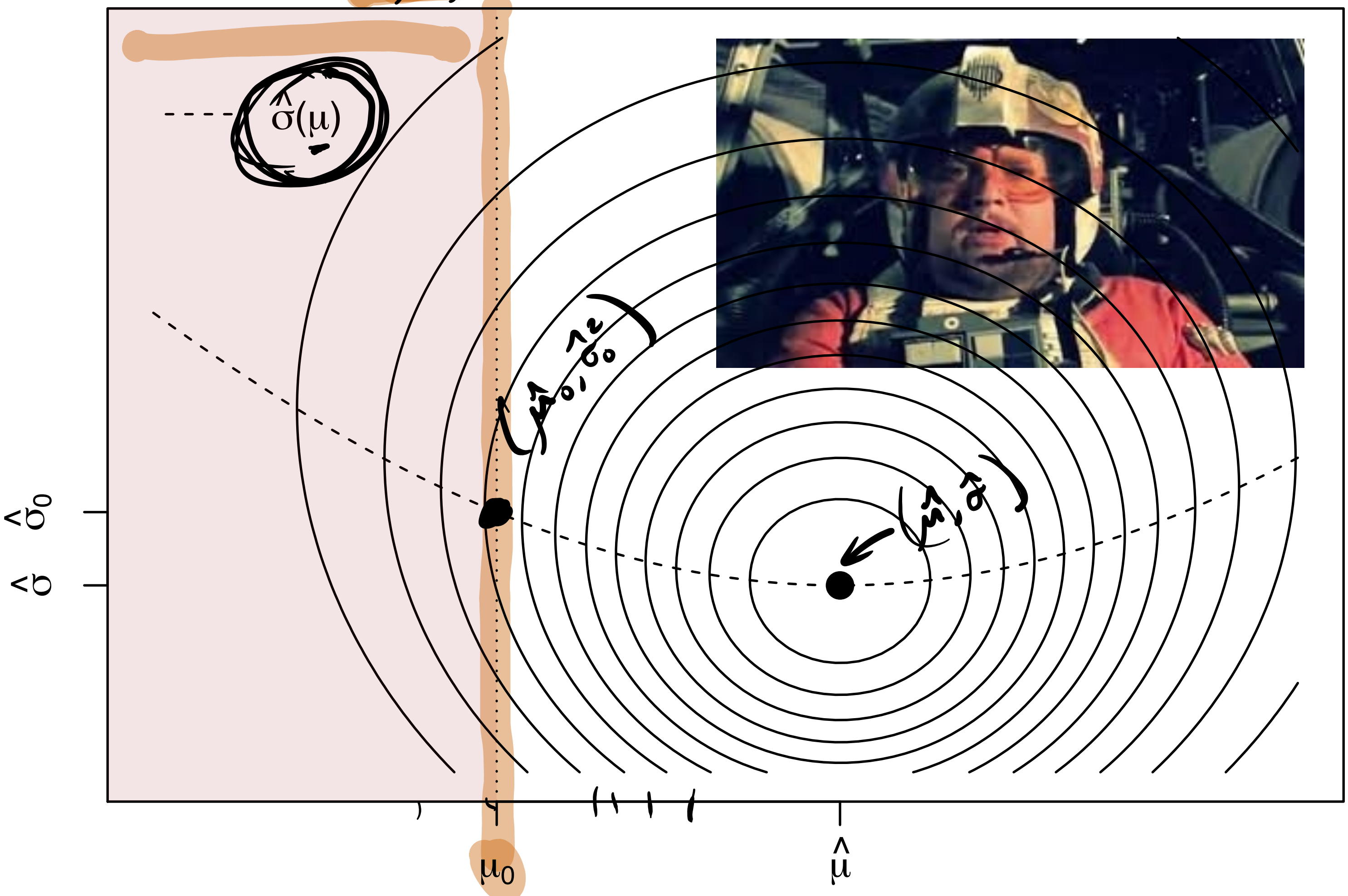
$$\beta(\mu) = P_\mu\left(\frac{\sqrt{n}\,(\bar{X}_n - \mu_0)}{S_n} > c \right)$$

$$\alpha = \text{size} = \sup_{\mu \le \mu_0} \beta(\mu) = \beta(\mu_0) = P_{\mu_0}\left(T > c \right),$$

when $T \sim t_{n-1}$,

So take $c = t_{n-1,\alpha}$.

$$(\hat{\mu}_0, \hat{\sigma}_0^2) = \underset{\{\mu \le \mu_0,\ \sigma^2 > 0\}}{\operatorname{argmax}} \ \ell(\mu, \sigma^2; \underset{\sim}{x})$$



Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, μ and σ^2 unknown, and consider

$$H_0\colon \sigma^2 = \sigma_0^2 \text{ versus } H_1\colon \sigma^2 \neq \sigma_0^2.$$

① Find LRT.

② Calibrate to have size α.

[redacted]

Steps:

[redacted]

$H_0: \sigma^2 = \sigma_0^2$ vs $H_1: \sigma^2 \neq \sigma_0^2$

$$LR(\underset{\sim}{X}) = \frac{\sup_{\mu \in \mathbb{R}, \sigma^2 = \sigma_0^2} L(\mu, \sigma^2; \underset{\sim}{X})}{\sup_{\mu \in \mathbb{R}, \sigma^2 > 0} L(\mu, \sigma^2; \underset{\sim}{X})} = \frac{L(\hat{\mu}_0, \overset{\sigma_0^2}{\hat{\sigma}_0^2}; \underset{\sim}{X})}{L(\hat{\mu}, \hat{\sigma}^2; \underset{\sim}{X})}$$

$$\begin{aligned}
\hat{\mu}_0 &= \underset{\mu \in \mathbb{R}}{\arg\max}\ L(\mu, \sigma_0^2; \underset{\sim}{X}) \\
&= \underset{\mu \in \mathbb{R}}{\arg\max}\ \ell(\mu, \sigma_0^2; \underset{\sim}{X}) \\
&= \underset{\mu \in \mathbb{R}}{\arg\max} \left\{ -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log \sigma_0^2 - \frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma_0^2} \right\} \\
&= \bar{x}_n
\end{aligned}$$

$$LR(\underset{\sim}{X}) = \frac{L(\bar{x}_n, \sigma_0^2; \underset{\sim}{X})}{L(\bar{x}_n, \hat{\sigma}_n^2; \underset{\sim}{X})} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma_0}\right)^n \exp\left[-\frac{\overbrace{\sum_{i=1}^n (x_i - \bar{x}_n)^2}^{n\hat{\sigma}_n^2}}{2\sigma_0^2}\right]}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\hat{\sigma}_n}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i - \bar{x}_n)^2}{2\underbrace{\hat{\sigma}_n^2}_{\frac{1}{n}\sum (x_i - \bar{x}_n)^2}}\right]}$$

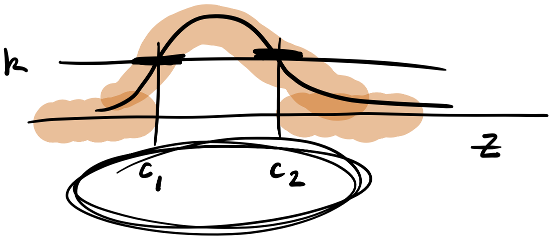
$$= \left(\frac{\hat{\sigma}_n}{\sigma_0}\right)^n \exp\left[-\frac{n\hat{\sigma}_n^2}{2\sigma_0^2}\right] e^{\frac{n}{2}}$$

$$= \left[\left(\frac{\hat{\sigma}_n^2}{\sigma_0^2}\right) \exp\left[-\frac{\hat{\sigma}_n^2}{\sigma_0^2}\right] e^{1} \right]^{n/2}$$

So LRT rejects H_0 when

$$\boxed{\left[\left(\frac{\hat{\sigma}_n^2}{\sigma_0^2}\right) \exp\left[-\frac{\hat{\sigma}_n^2}{\sigma_0^2}\right] e^{1} \right]^{n/2}} < k .$$

(2) $z e^{-z}$



LRT is equiv. to the test with rejection rule

$$\frac{\hat{\sigma}_n^2}{\sigma_0^2} < c_1 \quad \text{or} \quad \frac{\hat{\sigma}_n^2}{\sigma_0^2} > c_2$$

$$\Leftrightarrow \quad \frac{(n-1)S_n^2}{\sigma_0^2} < \underbrace{n c_1}_{c_1^*} \quad \text{or} \quad \frac{(n-1)S_n^2}{\sigma_0^2} > \underbrace{n c_2}_{c_2^*}$$

for some $c_1 < c_2$.

Note: $$\frac{\hat{\sigma}_n^2}{\sigma_0^2} = \frac{\frac{1}{n}\overbrace{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}^{(n-1)S_n^2}}{\sigma_0^2} = \frac{1}{n} \cdot \underbrace{\frac{(n-1)S_n^2}{\sigma_0^2}}_{\sim \chi^2_{n-1} \text{ if } \sigma^2 = \sigma_0^2.}$$

Can find $c_1^* < c_2^*$ by solving

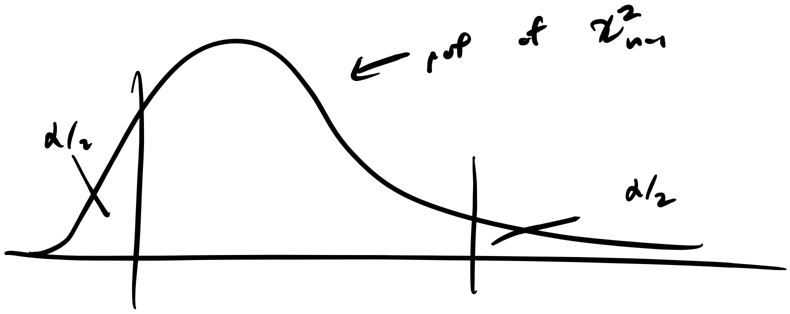
$$P(W < c_1^*) + P(W > c_2^*) = \alpha \,, \quad W \sim \chi^2_{n-1}$$

and

$$c_1^* e^{-c_1^*} = c_2^* e^{-c_2^*} .$$

Easier option: "Equal tails"

$$c_1^* = \chi^2_{n-1,\,1-\alpha/2} \qquad c_2^* = \chi^2_{n-1,\,\alpha/2}$$



$\chi^2_{n-1,\,1-\alpha/2} = \tilde{c}_1^*$ $\qquad$ $\chi^2_{n-1,\,\alpha/2} = \tilde{c}_2^*$

Wilk's Theorem

Let $\mathbf{X}_n$ be a random sample with likelihood $\mathcal{L}(\theta; \mathbf{X}_n)$, $\theta \in \Theta$, and consider

$$\boxed{H_0\colon \tau(\theta) = \tau_0 \text{ versus } H_1\colon \tau(\theta) \neq \tau_0,}$$

where $\dim(\Theta) = d$ and $\Theta_0 = \{\theta \in \Theta : \tau(\theta) = \tau_0\}$, with $\dim(\Theta_0) = d_0 < d$.

Then under H_0, we have

$$-2 \log \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; \mathbf{X}_n)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta; \mathbf{X}_n)} \xrightarrow{\text{D}} \chi^2_{d-d_0}$$

as $n \to \infty$, provided the conditions for asymptotic Normality of MLEs hold.

Exercise: Go through heuristics of proof for simple case $H_0\colon \theta = \theta_0$.

Wilks's theorem:

Assume the ML regularity conditions and consider testing

$$H_0: \tau(\theta) = \tau_0 \quad \text{vs} \quad H_1: \tau(\theta) \neq \tau_0 .$$

Let $d_0 = \dim(\Theta_0)$, $\Theta_0 = \{\theta \in \Theta : \tau(\theta) = \tau_0\}$, and $d = \dim(\Theta)$.

Then under H_0,

$$-2 \log \frac{\sup_{\theta \in \Theta_0} L(\theta; \underset{\sim}{X})}{\sup_{\theta \in \Theta} L(\theta; \underset{\sim}{X})} \xrightarrow{D} \chi^2_{d-d_0}$$

as $n \to \infty$.

Sketch of proof when $H_0: \theta = \theta_0$:

Firstly, under $H_0: \theta = \theta_0$, $d_0 = 0$, since $\dim(\{\theta_0\}) = 0$.

Assume $\theta = \theta_0$, i.e. the null is true. Then

$$-2 \log \frac{\sup_{\theta \in \Theta_0} L(\theta; \underset{\sim}{X})}{\sup_{\theta \in \Theta} L(\theta; \underset{\sim}{X})} = -2 \log \frac{L(\theta_0; \underset{\sim}{X})}{L(\hat{\theta}_n; \underset{\sim}{X})}, \quad \hat{\theta}_n \text{ the mle.}$$

Use expansion

$\ell(\theta_0; \underset{\sim}{X}) \approx \ell(\hat{\theta}_n; \underset{\sim}{X}) + s(\hat{\theta}_n; \underset{\sim}{X})^T (\theta_0 - \hat{\theta}_n) + \frac{1}{2} (\theta_0 - \hat{\theta}_n)^T \mathcal{H}(\hat{\theta}_n; \underset{\sim}{X}) (\theta_0 - \hat{\theta}_n)$

$$= -2 \left[\ell(\theta_0; \underset{\sim}{X}) - \ell(\hat{\theta}_n; \underset{\sim}{X}) \right]$$

$$\approx -2 \left[\underbrace{[s(\hat{\theta}_n; \underset{\sim}{X})]^T}_{=0} (\theta_0 - \hat{\theta}_n) + \frac{1}{2} (\theta_0 - \hat{\theta}_n)^T \mathcal{H}(\hat{\theta}_n; \underset{\sim}{X}) (\theta_0 - \hat{\theta}_n) \right]$$

$$= (\hat{\theta}_n - \theta_0)^T \left[-\mathcal{H}(\hat{\theta}_n; \underset{\sim}{X}) \right] (\hat{\theta}_n - \theta_0)$$

$-\mathcal{H}(\hat{\theta}_n; \underset{\sim}{X})$ is the observed Fisher information — the MLE of $I_n(\theta_0) = n I_1(\theta_0)$.

$$\approx n (\hat{\theta}_n - \theta_0)^T I_1(\theta_0) (\hat{\theta}_n - \theta_0)$$

since $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N(0, I_1^{-1}(\theta_0))$

$$\xrightarrow{D} \chi^2_d$$

Wilk's Theorem allows us to define the following test:

Definition (Asymptotic likelihood ratio test)

Under the settings of Wilk's theorem, the test with rejection rule

$$\text{Reject } H_0 \text{ iff } -2\log\frac{\sup_{\theta\in\Theta_0}\mathcal{L}(\theta;\mathbf{X}_n)}{\sup_{\theta\in\Theta}\mathcal{L}(\theta;\mathbf{X}_n)} > \chi^2_{d-d_0,\alpha}.$$

has size approximately α for large n. It is called the *asymptotic* LRT.

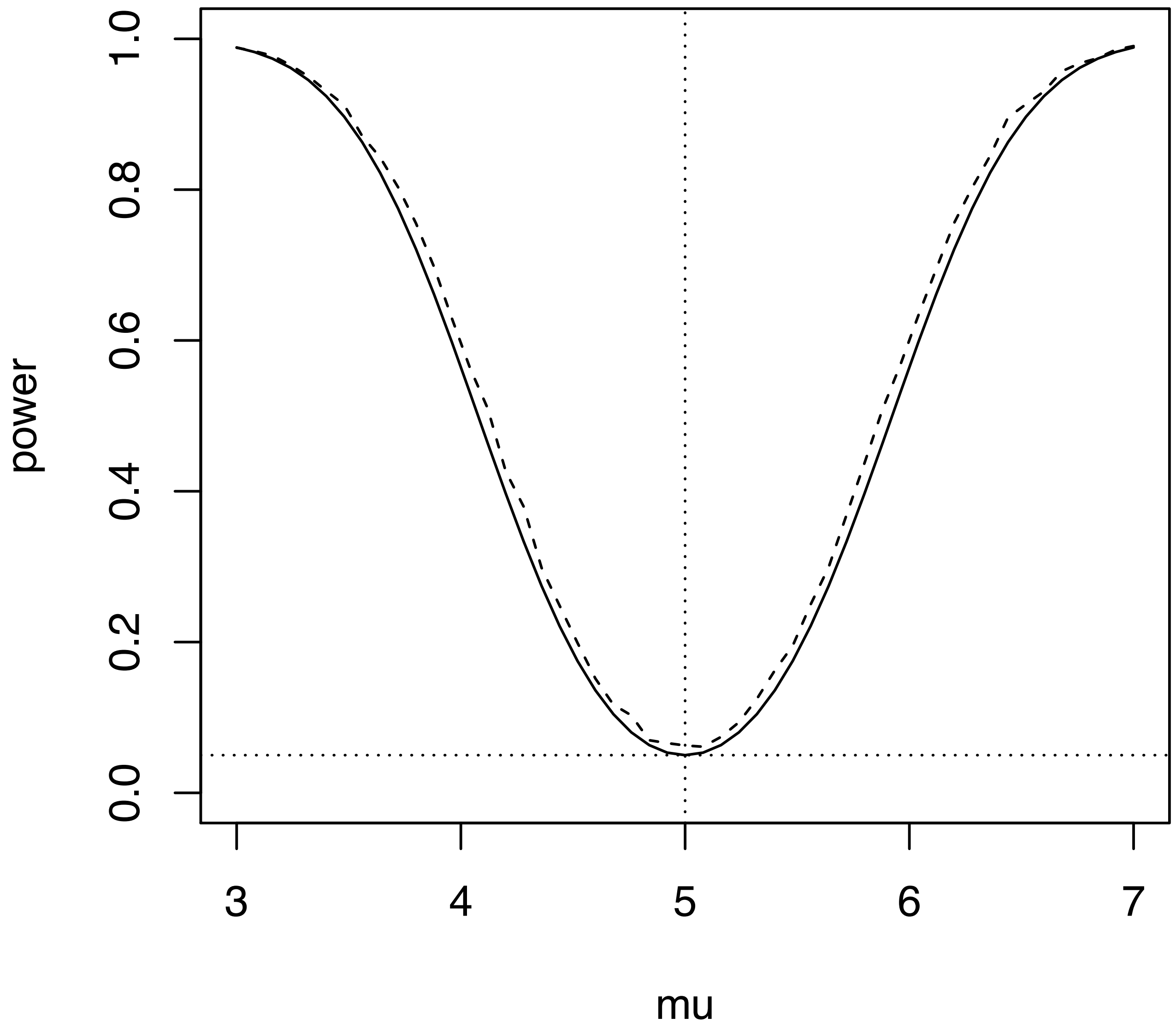
The asymptotic LRT is only for two-sided hypotheses.

- $d_0 = \dim(\Theta_0)$ is the number of parameters left unspecified by H_0.
- $d = \dim(\Theta)$ is the total number of unknown parameters.

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, μ and σ^2 unknown, and consider

$$H_0\colon \mu = \mu_0 \text{ versus } H_1\colon \mu \neq \mu_0.$$

1. Find the size-α asymptotic LRT.
2. Compare via simulation the power curves of the asymptotic LRT and the (non-asymptotic) LRT. Use the settings $n = 20$ and $\alpha = 0.05$.


power
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
mu

$p_X(x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$, $x = 0, 1, 2, \ldots$

$\lambda \in \underbrace{(0, \infty)}_{\Theta}$

$\lambda > 0$

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Poisson}(\lambda)$, λ unknown, and consider testing

$$H_0\colon \lambda = \lambda_0 \text{ versus } H_1\colon \lambda \neq \lambda_0.$$

$\dim(\Theta) = 1 = d$

$\Theta_0 = \{\lambda_0\}$ $\dim(\Theta_0) = 0 = d_0$

1. Give the decision rule of the asymptotic LRT.
2. Let $\lambda_0 = 3$ and run a simulation to get power curves for the test under the sample sizes $n = 5, 10, 20, 40$ using $\alpha = 0.05$.
3. Find the p-value of the asymptotic likelihood ratio test of

$$H_0\colon \lambda = 3 \text{ versus } H_1\colon \lambda \neq 3$$

associated with a sample of size $n = 25$ with sample mean equal to 3.5.

(1) $L(\lambda; \underset{\sim}{X}) = \dfrac{e^{-n\lambda}\lambda^{\Sigma X_i}}{\pi X_i!} = e^{-n\lambda}\lambda^{n\bar{X}_n} \Big/ \prod_{i=1}^{n} X_i!$

$$LR(\underset{\sim}{X}) = \frac{\sup_{\lambda \in \{\lambda_0\}} L(\lambda; \underset{\sim}{X})}{\sup_{\lambda > 0} L(\lambda; \underset{\sim}{X})} = \frac{L(\lambda_0; \underset{\sim}{X})}{L(\hat{\lambda}_n; \underset{\sim}{X})} \qquad \hat{\lambda}_n = \bar{X}_n$$

$$LR(\underset{\sim}{X}) = \frac{e^{-n\lambda_0} \lambda_0^{n\bar{X}_n} \Big/ \prod_{i=1}^{n} x_i!}{e^{-n\bar{X}_n} \bar{X}_n^{n\bar{X}_n} \Big/ \prod_{i=1}^{n} x_i!} = e^{-n(\lambda_0 - \bar{X}_n)} \left(\frac{\lambda_0}{\bar{X}_n}\right)^{n\bar{X}_n}$$

non. asympt. LRT rejects $H_0: \lambda = \lambda_0$ when $e^{-n(\lambda_0 - \bar{X}_n)} \left(\frac{\lambda_0}{\bar{X}_n}\right)^{n\bar{X}_n} < k$

$\iff \bar{X}_n < c_1$ or $\bar{X}_n > c_2$ for some c_1 and c_2.

<u>Asymp LRT</u>:

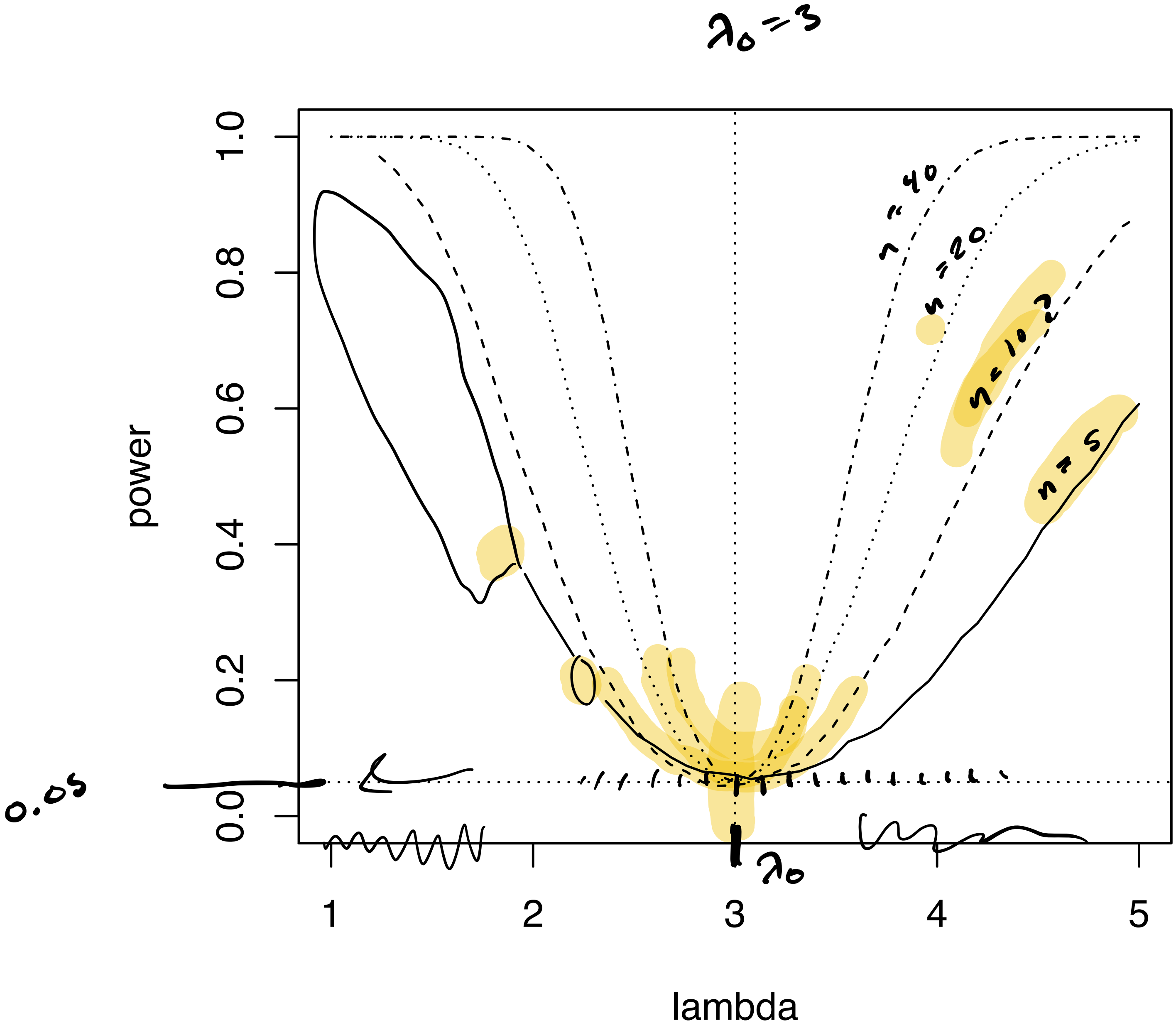
$$-2 \log LR(\underset{\sim}{X}) = -2 \log\left(e^{n(\lambda_0 - \bar{X}_n)} \left(\frac{\lambda_0}{\bar{X}_n}\right)^{n\bar{X}_n} \right)$$

$$= -2 \left[n(\lambda_0 - \bar{X}_n) + n\bar{X}_n \log\left(\frac{\lambda_0}{\bar{X}_n}\right) \right]$$

If truly $\lambda = \lambda_0$, then

$$-2 \left[n(\lambda_0 - \bar{X}_n) + n\bar{X}_n \log\left(\frac{\lambda_0}{\bar{X}_n}\right) \right] \xrightarrow{D} \chi^2_2. \qquad n=5$$

So ALRT rejects $H_0: \lambda = \lambda_0$ if $-2 \left[n(\lambda_0 - \bar{X}_n) + n\bar{X}_n \log\left(\frac{\lambda_0}{\bar{X}_n}\right) \right] > \chi^2_{1,\alpha}$

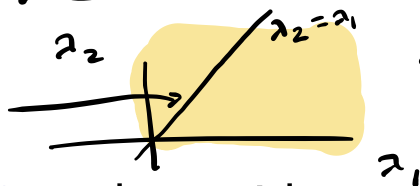

$\lambda_0 = 3$
power
1.0
0.8
0.6
0.4
0.2
0.0
0.05
n = 40
n = 20
n = 10
n = 5
λ_0
1
2
3
4
5
lambda

$X_{11},\dots,X_{1n_1} \overset{ind}{\sim}$ Exponential (λ_1)

$X_{21},\dots,X_{2n_2} \overset{ind}{\sim}$ Exponential (λ_2)

$\Theta = \{(\lambda_1,\lambda_2): \lambda_1>0,\ \lambda_2>0\}$

Θ_0



$\mathbb{R}^2$

Exercise: Let $X_{k1},\dots,X_{kn_k} \overset{\text{ind}}{\sim} \text{Exponential}(\lambda_k)$, $k=1,2$ and consider

$\tau(\theta)=\tau_0$ $\quad$ $\tau(\theta)\neq\tau_0$

$$\boxed{H_0\colon \lambda_1=\lambda_2} \text{ versus } H_1\colon \lambda_1\neq\lambda_2.$$

$\dim(\Theta) = 2 = d$

$\Theta_0 : \{(\lambda_1,\lambda_2): \lambda_1=\lambda_2,\ \lambda_1>0\}$ $\quad$ $\dim(\Theta_0) = 1 = d_0$

- Give the decision rule of the asymptotic LRT.
- Plot power curves of the test with $\alpha = 0.05$ under

$$(n_1,n_2) = (10,20),(20,40),(50,100),(100,200)$$

  when $\lambda_1 = 3$ with λ_2 varying from 1 to 5.
- Find the p-value of the asymptotic LRT associated with observing $\bar{X}_1 = 4.1$ and $\bar{X}_2 = 3.2$ when $n_1 = 30$ and $n_2 = 34$.

$f_X(x;\lambda_1) = \frac{1}{\lambda_1} e^{-x/\lambda_1}\, \mathbb{1}(x>0)$ $\qquad$ $\theta = (\lambda_1,\lambda_2)$

$H_0: \lambda_1=\lambda_2 \iff \lambda_1-\lambda_2=0,$ $\quad$ $\tau(\theta)=0,\ \tau(\lambda_1,\lambda_2)=\lambda_1-\lambda_2.$

$$L(\lambda_1,\lambda_2;\underset{\sim}{X}_1,\underset{\sim}{X}_2) = \prod_{i=1}^{n_1} \frac{1}{\lambda_1} e^{-x_{1i}/\lambda_1} \cdot \prod_{j=1}^{n_2} \frac{1}{\lambda_2} e^{-x_{2j}/\lambda_2}$$

$$= \left(\frac{1}{\lambda_1}\right)^{n_1} e^{-\frac{n_1\bar{x}_1}{\lambda_1}} \left(\frac{1}{\lambda_2}\right)^{n_2} e^{-\frac{n_2\bar{x}_2}{\lambda_2}}$$

$$LR(\underset{\sim}{X}_1,\underset{\sim}{X}_2) = \frac{\sup_{\lambda_1=\lambda_2,\,\lambda_1>0} L(\lambda_1,\lambda_2;\underset{\sim}{X}_1,\underset{\sim}{X}_2)}{\sup_{\lambda_1>0,\,\lambda_2>0} L(\lambda_1,\lambda_2;\underset{\sim}{X}_1,\underset{\sim}{X}_2)}$$

$$= \frac{L(\hat{\lambda}_0,\hat{\lambda}_0;\underset{\sim}{X}_1,\underset{\sim}{X}_2)}{L(\bar{x}_1,\bar{x}_2;\underset{\sim}{X}_1,\underset{\sim}{X}_2)},$$

when $\hat{\lambda}_0 = \underset{\lambda}{\operatorname{argmax}}\ L(\lambda,\lambda;\underset{\sim}{X}_1,\underset{\sim}{X}_2)$

$$\ell(\lambda,\lambda;\underset{\sim}{X}_1,\underset{\sim}{X}_2) = \log\left[\left(\frac{1}{\lambda}\right)^{n_1} e^{-\frac{n_1\bar{x}_1}{\lambda}} \left(\frac{1}{\lambda}\right)^{n_2} e^{-\frac{n_2\bar{x}_2}{\lambda}}\right]$$

$$= -(n_1+n_2)\log\lambda - \frac{(n_1\bar{x}_1 + n_2\bar{x}_2)}{\lambda}$$

$$\frac{\partial}{\partial\lambda}\ell(\lambda,\lambda;\underset{\sim}{X}_1,\underset{\sim}{X}_2) = \frac{-(n_1+n_2)}{\lambda} + \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{\lambda^2} \overset{\text{set}}{=} 0$$

$$\hat{\lambda}_0 = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2} = \frac{\sum_{i=1}^{n_1} X_{1i} + \sum_{j=1}^{n_2} X_{2j}}{n_1+n_2}.$$

$$LR(\underset{\sim}{X}_1, \underset{\sim}{X}_2) = \frac{\left(\dfrac{1}{\frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}}\right)^{n_1} e^{-n_1\bar{x}_1\left(\frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}\right)^{-1}} \left(\dfrac{1}{\frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}}\right)^{n_2} e^{-n_2\bar{x}_2\left(\frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}\right)^{-1}}}{\left(\frac{1}{\bar{x}_1}\right)^{n_1} e^{-\frac{n_1\cancel{\bar{x}_1}}{\cancel{\bar{x}_1}}} \left(\frac{1}{\bar{x}_2}\right)^{n_2} e^{-\frac{n_2\cancel{\bar{x}_2}}{\cancel{\bar{x}_2}}}}$$

$$= \frac{\left(\frac{n_1+n_2}{n_1\bar{x}_1 + n_2\bar{x}_2}\right)^{n_1+n_2}}{\left(\frac{1}{\bar{x}_1}\right)^{n_1}\left(\frac{1}{\bar{x}_2}\right)^{n_2}} \frac{e^{-(n_1+n_2)}}{e^{-(n_1+n_2)}}$$

$$= \bar{x}_1^{n_1}\ \bar{x}_2^{n_2} \left(\frac{1}{n_1\bar{x}_1 + n_2\bar{x}_2}\right)^{n_1+n_2} (n_1+n_2)^{n_1+n_2}$$

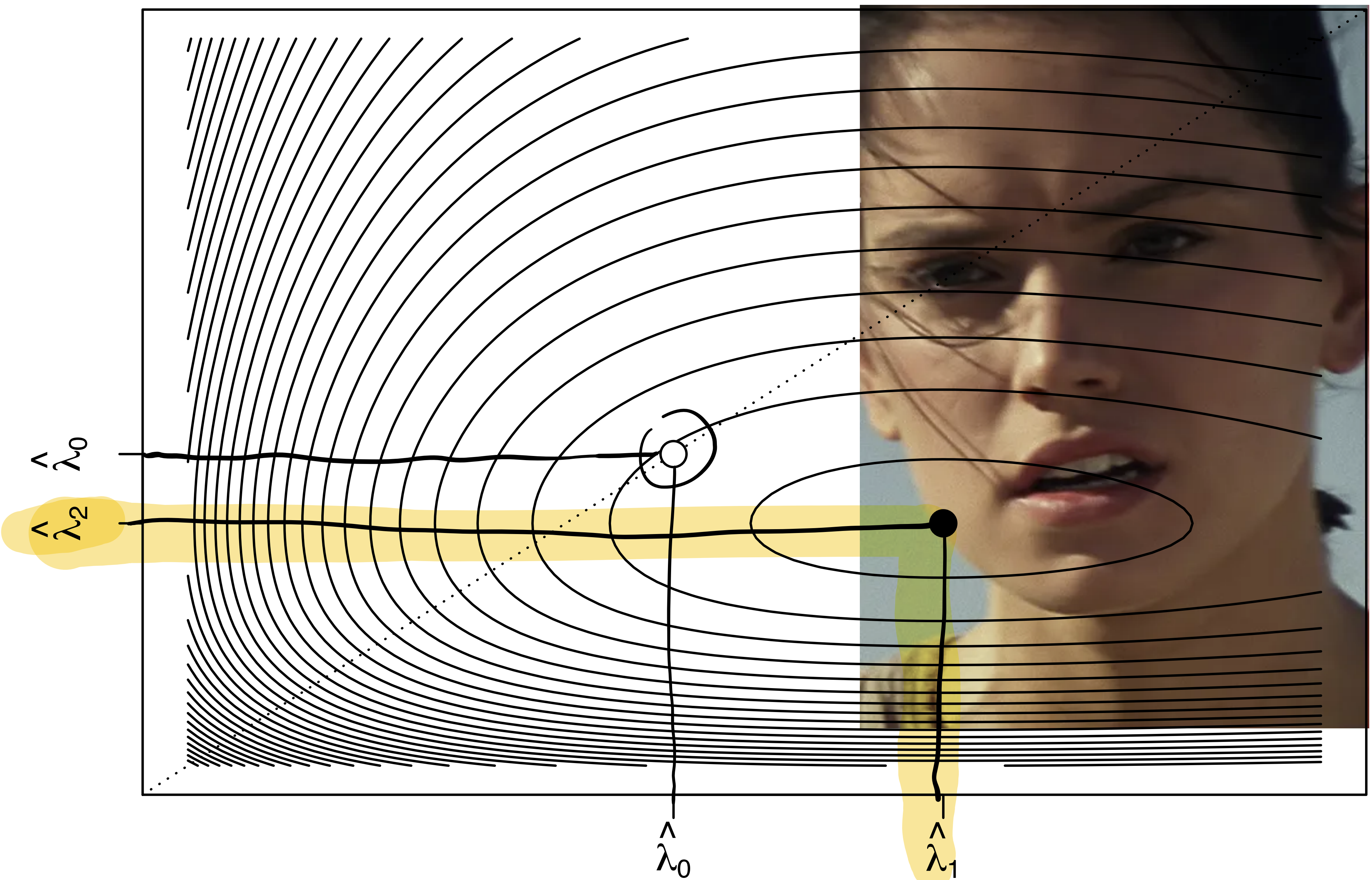
$$= \left(\frac{n_1\bar{x}_1}{n_1\bar{x}_1 + n_2\bar{x}_2}\right)^{n_1} \left(\frac{n_2\bar{x}_2}{n_1\bar{x}_1 + n_2\bar{x}_2}\right)^{n_2} \left(\frac{n_1+n_2}{n_1}\right)^{n_1} \left(\frac{n_1+n_2}{n_2}\right)^{n_2}$$

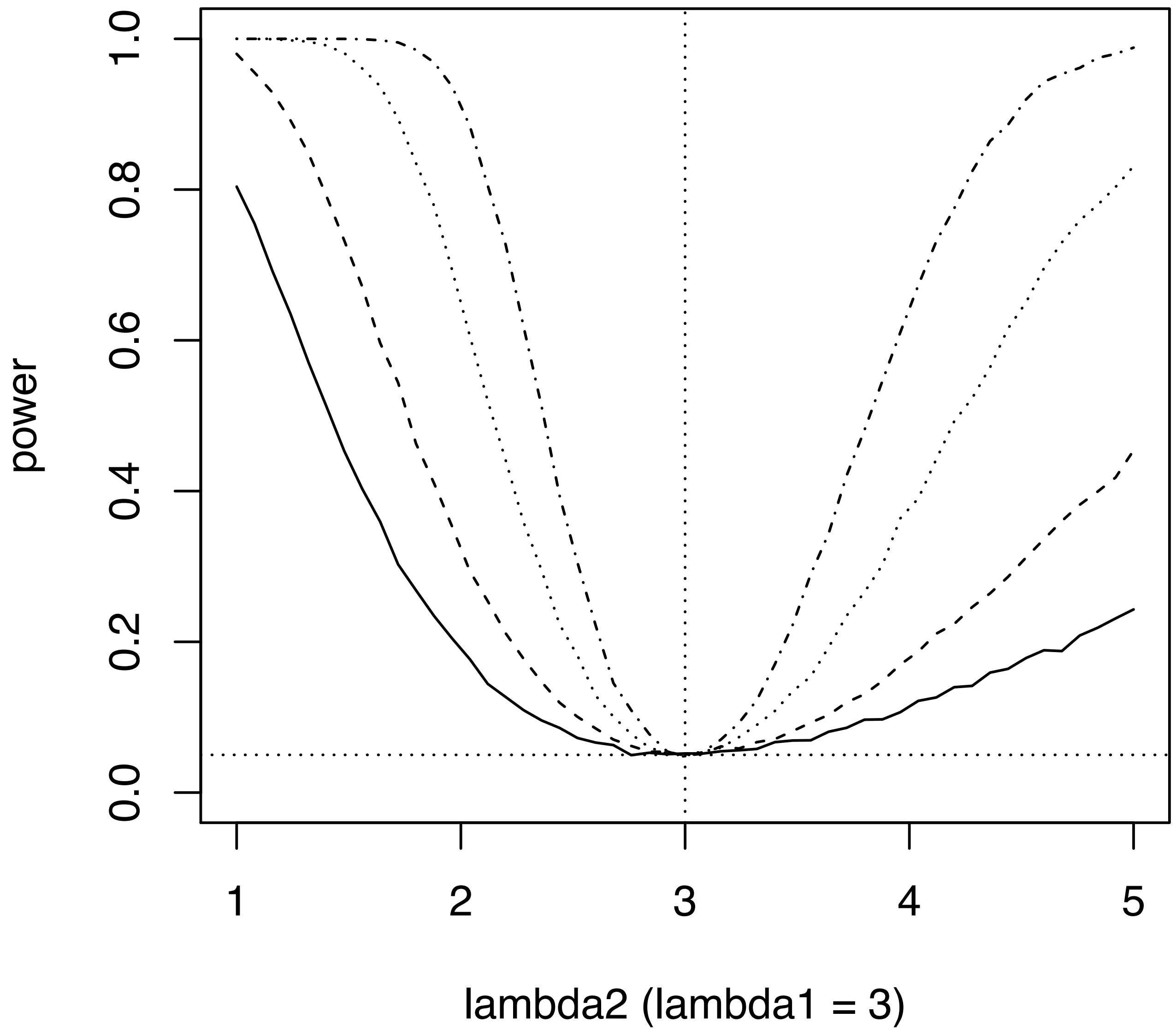
$$-2\log LR(\underset{\sim}{X}_1, \underset{\sim}{X}_2) = -2\log\left(\frac{\left(\frac{n_1+n_2}{n_1\bar{x}_1 + n_2\bar{x}_2}\right)^{n_1+n_2}}{\left(\frac{1}{\bar{x}_1}\right)^{n_1}\left(\frac{1}{\bar{x}_2}\right)^{n_2}}\right)$$

$$= -2\left[(n_1+n_2)\log\left(\frac{n_1+n_2}{n_1\bar{x}_1 + n_2\bar{x}_2}\right) - n_1\log\left(\frac{1}{\bar{x}_1}\right) - n_2\log\left(\frac{1}{\bar{x}_2}\right)\right]$$

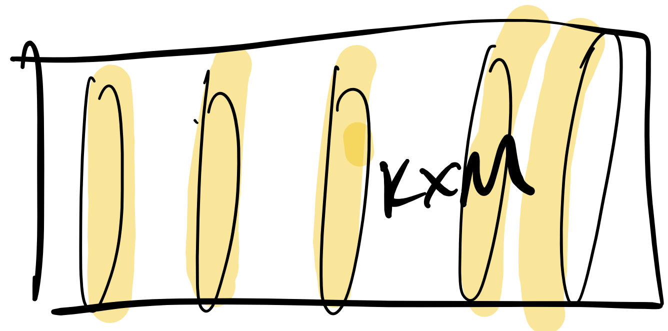
Reject H_0 when this $> \chi^2_{1,\alpha}$

$L(\lambda_1, \lambda_2; \underset{\sim}{X}_1, \underset{\sim}{X}_2)$





power
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
lambda2 (lambda1 = 3)



Exercise: For some $K \geq 1$ and $M \geq 1$, suppose we observe indep. random vectors

$$Y_m \sim \text{Multinomial}(n_m, p_{m1}, \dots, p_{mK}), \quad \text{for } m = 1, \dots, M.$$

Derive the asymptotic size-α likelihood ratio test for

$$H_0 : p_{ik} = p_{jk} \text{ for all } i \neq j \text{ for each } k = 1, \dots, K.$$

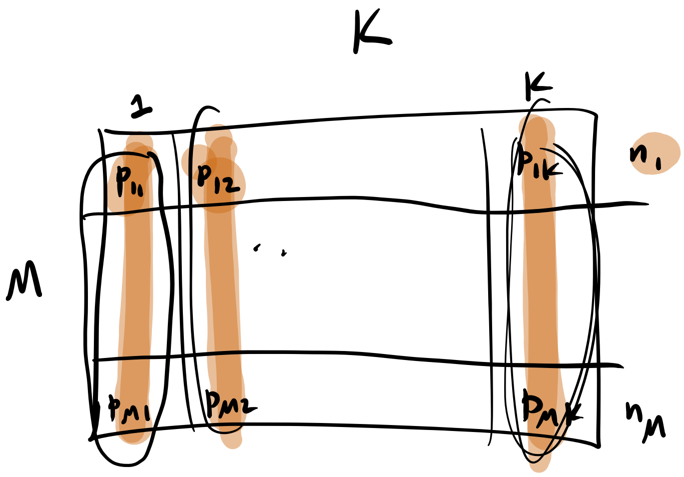
$$d - d_0 = (K-1)(M-1)$$

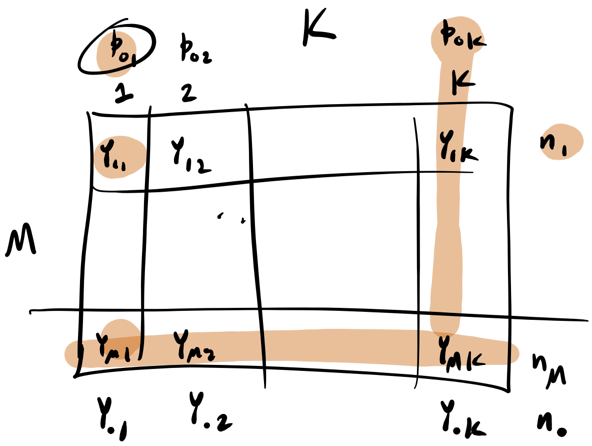
<u>**EXERCISE**</u>: For some $K \geq 1$, $M \geq 1$, suppose

$$Y_m \sim \text{Multinomial}(n_m, p_{m1}, \dots, p_{mK}), \quad m = 1, \dots, M$$

are indep. Derive size-α ALRT for

$$H_0: \underline{p_{ik} = p_{jk}} \quad \forall\, i \neq j \quad \text{for each } k = 1, \dots, K.$$





$$p_{11} = \dots = p_{M1} = p_{01}, \quad \dots, \quad p_{1K} = \dots = p_{MK} = p_{0K}$$

<u>**Remember:**</u>

$$\underset{K \times 1}{\underset{\sim}{X}} \sim \text{Multinomial}(n, p_1, \dots, p_K)$$

parameter space

$$\{ p_1, \dots, p_K \in [0,1] : p_1 + \dots + p_K = 1 \}$$

has dim $K-1$. (Simplex)

$$p(\underset{\sim}{x}; p_1, \dots, p_K) = \left(\frac{n!}{x_1! \cdot \ldots \cdot x_K!} \right) p_1^{x_1} \cdot \ldots \cdot p_K^{x_K}$$

$$p_K = 1 - (p_1 + \dots + p_{K-1})$$

$$p_1 + \dots + p_K = 1 \qquad x_1 + \dots + x_K = n$$

<u>**MLEs of $p_1, \dots, p_K$:**</u>

$$L(p_1, \dots, p_K; \underset{\sim}{x}) = \left(\frac{n!}{x_1! \cdot \ldots \cdot x_K!} \right) p_1^{x_1} \cdot \ldots \cdot p_K^{x_K}$$

$$\ell(p_1,\dots,p_K ; \underset{\sim}{X}) = \log\left(\quad\right) + x_1 \log p_1 + \dots + x_K \log p_K$$

Lagrange Method:

Maximize $\quad x_1 \log p_1 + \dots + x_K \log p_K$

subject to $\quad p_1 + \dots + p_K = 1.$

1. Make new function

$$\mathcal{L}(p_1,\dots,p_K,\lambda) = x_1 \log p_1 + \dots + x_K \log p_K + \underset{\sim}{\lambda}\left(1 - \sum_{k=1}^{K} p_k\right)$$

2. Take deriv w.r.t. $p_1,\dots,p_K$ and λ, set **ALL** $=0$ and solve system.

$$\left\{\begin{aligned} \frac{\partial}{\partial p_1}\mathcal{L}(\cdot\cdot) &= \frac{x_1}{p_1} - \lambda = 0 \\ &\vdots \\ \frac{\partial}{\partial p_K}\mathcal{L}(\cdot\cdot) &= \frac{x_K}{p_K} - \lambda = 0 \\ \frac{\partial}{\partial \lambda}\mathcal{L}(\cdot) &= 1 - \sum_{k=1}^{K} p_k = 0 \end{aligned}\right\}$$

$K+1$ unknowns

$K+1$ equations

$$\left.\begin{aligned} x_1 &= \lambda p_1 \\ &\vdots \\ x_K &= \lambda p_K \end{aligned}\right\} \Rightarrow \sum_{k=1}^{K} x_k = \sum_{k=1}^{K} \lambda p_k = \lambda \sum_{k=1}^{K} p_k \cdot 1$$

$$\Rightarrow \boxed{\lambda = n}$$

$$\left.\begin{array}{l} \dfrac{x_1}{p_1} - n = 0 \\ \vdots \\ \dfrac{x_k}{p_k} - n = 0 \end{array}\right) \Rightarrow \begin{array}{l} \hat{p}_1 = \dfrac{x_1}{n} \\ \vdots \\ \hat{p}_k = \dfrac{x_k}{n}. \end{array}$$

$$L\left(\{p_{mk};\ 1 \le m \le M,\ 1 \le k \le K\};\ \underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M\right)$$

$$= \prod_{m=1}^{M} \left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) p_{m1}^{Y_{m1}} \cdot \ldots \cdot p_{mK}^{Y_{mK}}$$

$$\ell\left(\{p_{mk};\ 1 \le m \le M,\ 1 \le k \le K\};\ \underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M\right)$$

$$= \sum_{m=1}^{M} \log\left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) + \sum_{m=1}^{M} \log\left(p_{m1}^{Y_{m1}} \cdot \ldots \cdot p_{mK}^{Y_{mK}}\right)$$

$$= \sum_{m=1}^{M} \log\left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) + \sum_{m=1}^{M} \sum_{k=1}^{K} Y_{mk} \log p_{mk}$$

MLEs are $\hat{p}_{mk} = \dfrac{Y_{mk}}{n_m}$

Restricted MLEs: Under H_0:

$p_{11} = \dots = p_{M1} = p_{01}, \dots, p_{1K} = \dots = p_{MK} = p_{0K}$

$$\ell(p_{01}, \dots, p_{0K}; \underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M)$$

$$= \sum_{m=1}^{M} \log\left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) + \sum_{k=1}^{K} \underbrace{\sum_{m=1}^{M} Y_{mk}}_{Y_{\cdot k}} \log p_{0k}$$

Maximize this, subject to $p_{01} + \dots + p_{0K} = 1$.

You get $\hat{p}_{0k} = \dfrac{Y_{\cdot k}}{n_{\cdot}}$

$$LR(\underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M) = \frac{\sup\limits_{\text{over null space}} L(\{p_{mk} : 1 \le m \le M, 1 \le k \le K\}; \underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M\}}{\sup\limits_{\text{over entire space}} L(\{p_{mk} : 1 \le m \le M, 1 \le k \le K\}; \underset{\sim}{Y}_1, \dots, \underset{\sim}{Y}_M\}}$$

$$= \frac{\prod_{m=1}^{M} \left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) \hat{p}_{01}^{Y_{m1}} \cdot \ldots \cdot \hat{p}_{0K}^{Y_{mK}}}{\prod_{m=1}^{M} \left(\frac{n_m!}{Y_{m1}! \cdot \ldots \cdot Y_{mK}!}\right) \hat{p}_{m1}^{Y_{m1}} \cdot \ldots \cdot \hat{p}_{mK}^{Y_{mK}}}$$

$$= \prod_{m=1}^{M} \left(\frac{\hat{p}_{01}}{\hat{p}_{m1}}\right)^{Y_{m1}} \cdot \ldots \cdot \left(\frac{\hat{p}_{0K}}{\hat{p}_{mK}}\right)^{Y_{mK}}$$

$$-2\log LR(\underset{\sim}{Y}_1,\dots,\underset{\sim}{Y}_M)$$

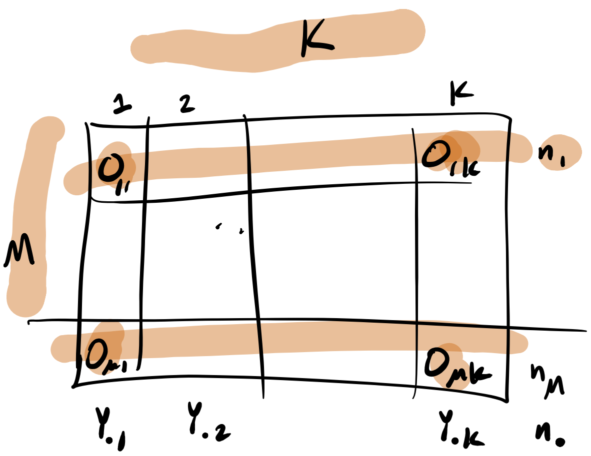
$$= -2\sum_{m=1}^{M}\left[Y_{m1}\log\left(\frac{\hat{p}_{01}}{\hat{p}_{m1}}\right)+\dots+Y_{mK}\log\left(\frac{\hat{p}_{0K}}{\hat{p}_{mK}}\right)\right]$$

$$= -2\sum_{m=1}^{M}\sum_{k=1}^{K} Y_{mk}\log\left(\frac{\hat{p}_{0k}}{\hat{p}_{mk}}\right)$$

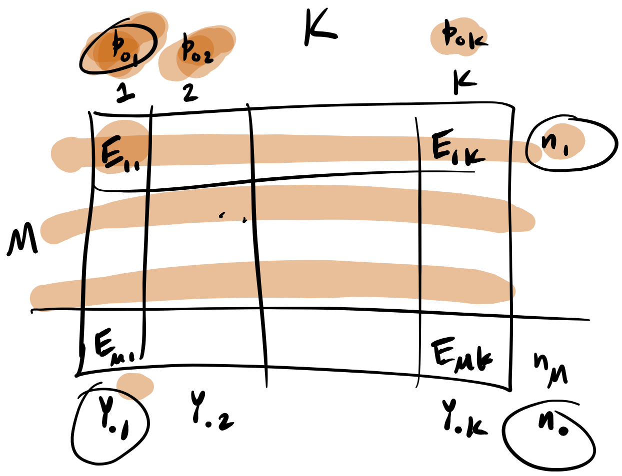
$$= 2\sum_{m=1}^{M}\sum_{k=1}^{K} Y_{mk}\log\left(\frac{\overbrace{n_m\cdot\hat{p}_{mk}}^{Y_{mk}}}{n_m\cdot\hat{p}_{0k}}\right)$$

$$= 2\sum_{m=1}^{M}\sum_{k=1}^{K} O_{mk}\log\left(\frac{O_{mk}}{E_{mk}}\right)$$ ← ALR test statistic.

Observed



Expected under H_0



$$E_{mk} = \frac{Y_{\cdot k}\cdot n_m}{n_\cdot} = \hat{p}_{0k}\cdot n_m$$

$$\frac{Y_{\cdot k}}{n_\cdot} = \hat{p}_{0k}$$

ALRT rejects H_0 when

$$2 \sum_{m=1}^{M} \sum_{k=1}^{K} O_{mk} \log\left(\frac{O_{mk}}{E_{mk}}\right) > \chi^2_{d-d_0,\,\alpha}$$

$$\approx$$

$$\sum_{m=1}^{M} \sum_{k=1}^{K} \frac{(O_{mk} - E_{mk})^2}{E_{mk}}$$

$\dim(\Theta_0) = K-1$

$\dim(\Theta) = M(K-1)$

$$d - d_0 = M(K-1) - (K-1) = \underline{\underline{(M-1)(K-1)}}$$