Hypothesis testing in non-standard situations

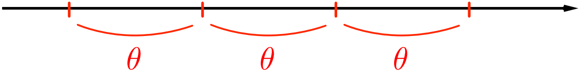
Davy Paindaveine

Université libre de Bruxelles

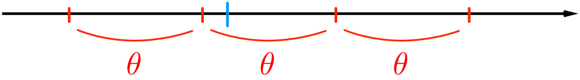
Godeaux Lecture, Brussels, May 2017


71 De Brouckère
9040
www.stib.be

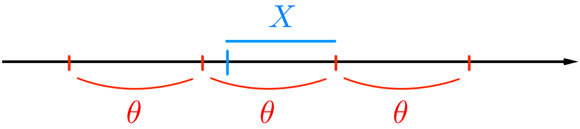

71 De Brouckère
9040
www.stib.be


θ
θ
θ

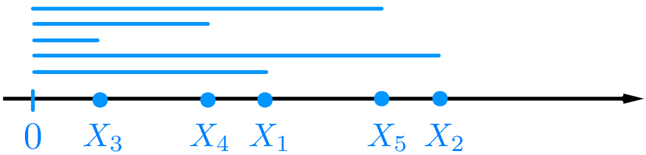

71 De Brouckère
9040
www.stib.be


θ
θ
θ


71 De Brouckère
9040
www.stib.be

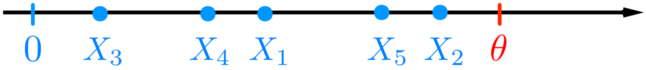

X
θ
θ
θ

Based on independent waiting times $X_1, \ldots, X_n$, how to estimate θ ?



Based on independent waiting times $X_1, \ldots, X_n$, how to estimate θ ?

Based on independent waiting times $X_1, \ldots, X_n$, how to estimate θ ?

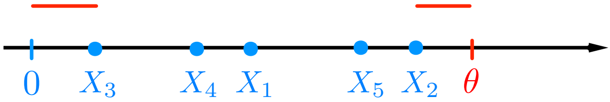


Possible estimators $\hat{\theta}$ include

- $2\bar{X}_n = 2(\frac{1}{n}\sum_{i=1}^{n} X_i)$
- $2\,\mathrm{Median}(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n) + \min(X_1, \ldots, X_n)$
- $\frac{n+1}{n}\max(X_1, \ldots, X_n)$
- $(n+1)\min(X_1, \ldots, X_n)$

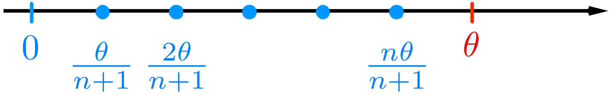
Possible estimators $\hat{\theta}$ include

- $2\bar{X}_n = 2(\frac{1}{n}\sum_{i=1}^{n} X_i)$
- $2\,\mathrm{Median}(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n) + \min(X_1, \ldots, X_n)$
- $\frac{n+1}{n}\max(X_1, \ldots, X_n)$
- $(n+1)\min(X_1, \ldots, X_n)$



Possible estimators $\hat{\theta}$ include

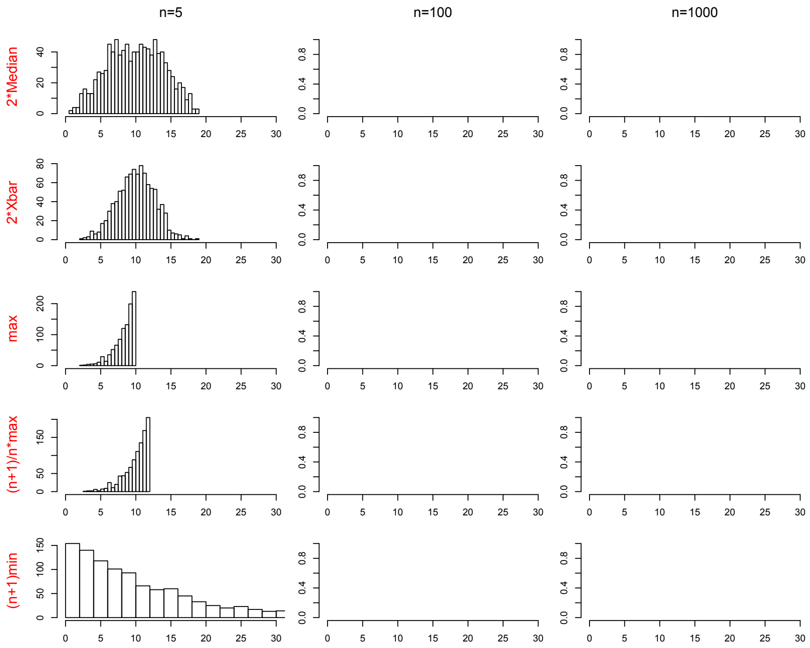
- $2\bar{X}_n = 2(\frac{1}{n}\sum_{i=1}^{n} X_i)$
- $2\,\mathrm{Median}(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n)$
- $\max(X_1, \ldots, X_n) + \min(X_1, \ldots, X_n)$
- $\frac{n+1}{n}\max(X_1, \ldots, X_n)$
- $(n+1)\min(X_1, \ldots, X_n)$



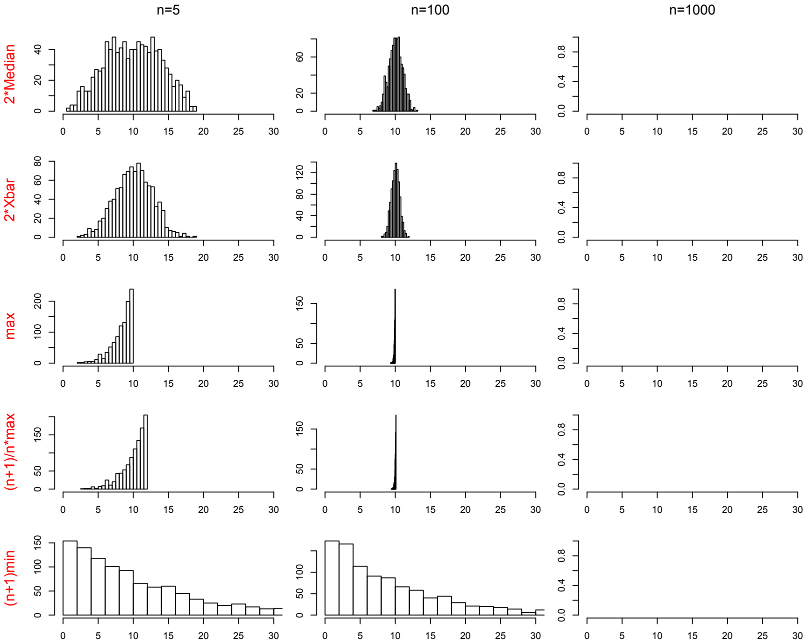
Possible estimators $\hat{\theta}$ include

one go for $n = 5$ and $\theta = 10$

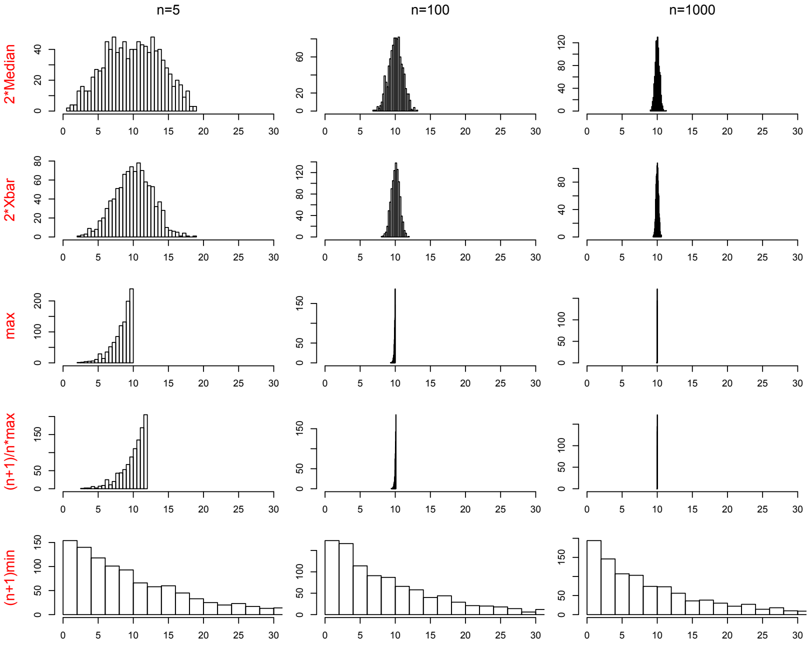
- $2\bar{X}_n = 2(\frac{1}{n}\sum_{i=1}^{n} X_i)$ ____ 10.29
- $2\,\mathrm{Median}(X_1, \ldots, X_n)$ ____ 11.47
- $\max(X_1, \ldots, X_n)$ ____ 9.44
- $\max(X_1, \ldots, X_n) + \min(X_1, \ldots, X_n)$ ____ 11.12
- $\frac{n+1}{n}\max(X_1, \ldots, X_n)$ ____ 11.33
- $(n+1)\min(X_1, \ldots, X_n)$ ____ 10.08



n=5
n=100
n=1000
2*Median
2*Xbar
max
(n+1)/n*max
(n+1)min



n=5
n=100
n=1000
2*Median
2*Xbar
max
(n+1)/n*max
(n+1)min



n=5
n=100
n=1000
2*Median
2*Xbar
max
(n+1)/n*max
(n+1)min

$$2\,\mathrm{Median}(X_1, \ldots, X_n) \xrightarrow{\mathcal{D}} \theta$$

$$2\bar{X}_n \xrightarrow{\mathcal{D}} \theta$$

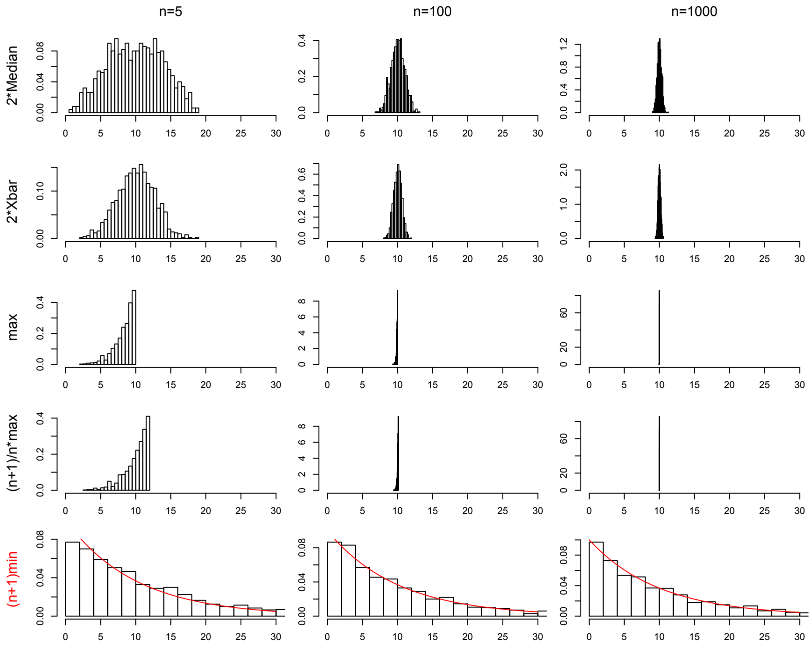
$$\frac{n+1}{n}\max(X_1, \ldots, X_n) \xrightarrow{\mathcal{D}} \theta$$

$$\max(X_1, \ldots, X_n) \xrightarrow{\mathcal{D}} \theta$$

and

$$(n+1)\min(X_1, \ldots, X_n) \xrightarrow{\mathcal{D}} Z,$$

where Z has density $(1/\theta)\exp(-z/\theta)$ on the positive halfline.



n=5
n=100
n=1000
2*Median
2*Xbar
max
(n+1)/n*max
(n+1)min

$$\sqrt{n}\,\big(\underbrace{2\,\mathrm{Median}(X_1,\ldots,X_n)-\theta}_{\xrightarrow{\mathcal{D}}0}\big)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\theta^2)$$

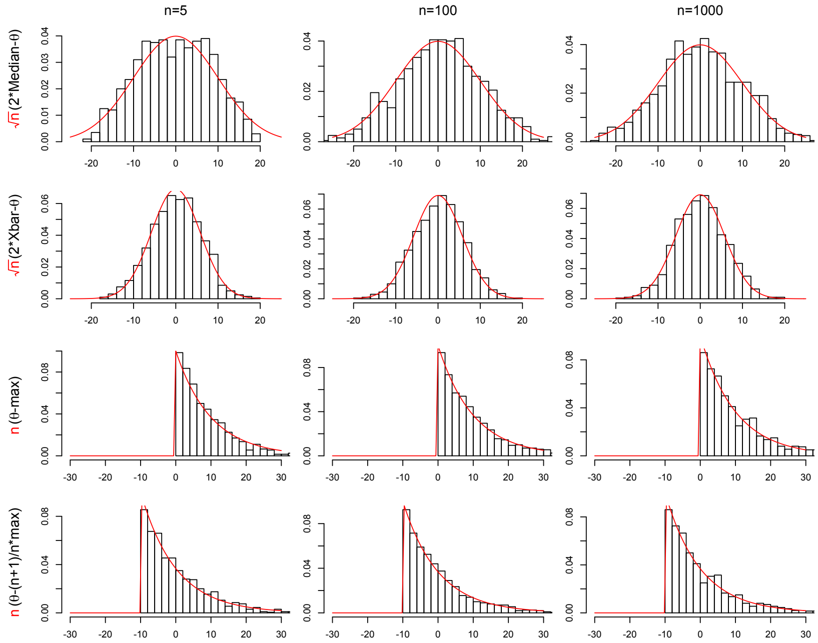
$$\sqrt{n}\,\big(\underbrace{2\bar{X}_n-\theta}_{\xrightarrow{\mathcal{D}}0}\big)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\theta^2/3)$$

$$n\,\big(\underbrace{\theta-\max(X_1,\ldots,X_n)}_{\xrightarrow{\mathcal{D}}0}\big)\xrightarrow{\mathcal{D}}Z$$

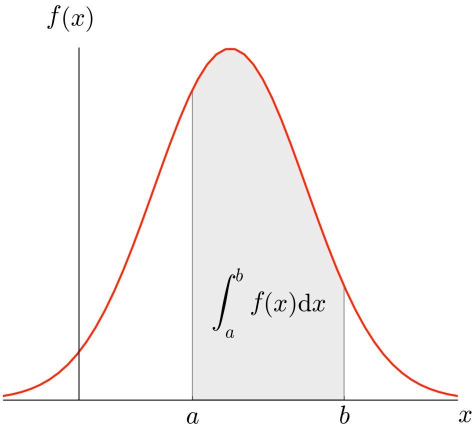
$$n\,\big(\underbrace{\theta-\tfrac{n+1}{n}\max(X_1,\ldots,X_n)}_{\xrightarrow{\mathcal{D}}0}\big)\xrightarrow{\mathcal{D}}Z-\theta,$$

with the same Z as above.

- Estimators may have different consistency rates (1st order)
- Estimators with a common rate may have different accuracies (2nd order)


n=5
n=100
n=1000
√n (2*Median-θ)
√n (2*Xbar-θ)
n (θ-max)
n (θ-(n+1)/n*max)

The amount of beer X is such that $P[a \leq X \leq b] = \int_a^b f(x)\,dx$, with f unknown



Our interest below is $\mu = \mathrm{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx$

- In estimation, you want to guess the value of μ
- In hypothesis testing, you only want to know whether (for a given μ_0)

$$\mathcal{H}_0 : \mu \geq \mu_0 \quad (\textit{the "null"})$$

or

$$\mathcal{H}_1 : \mu < \mu_0 \quad (\textit{the "alternative"})$$

(pick the μ_0 announced by the barman and sue him if $\mathcal{H}_1$ is shown to hold)

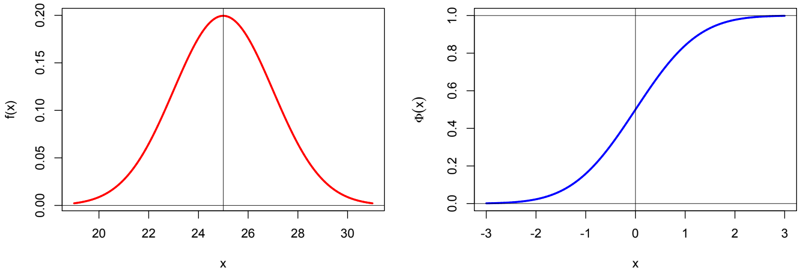
Based on a sample $X_1, \ldots, X_n$, the "test" associated with $B(\subset \mathbb{R}^n)$ rejects $\mathcal{H}_0$ in favour of $\mathcal{H}_1$ iff $(X_1, \ldots, X_n) \in B$.

Assume that $\sigma^2 = \mathrm{Var}[X_1] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx < \infty$. Fix $\alpha \in (0,1)$.

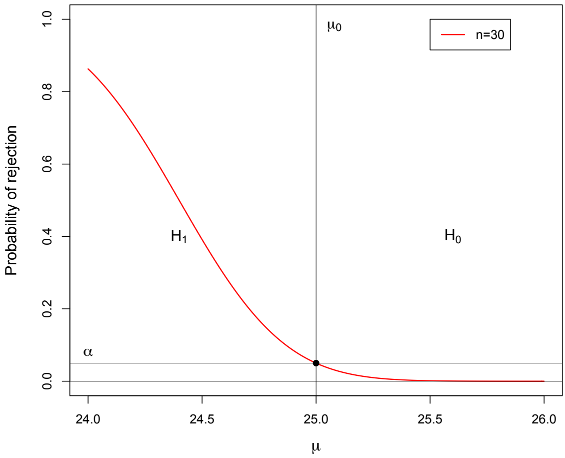
The textbook "level-α" test rejects $\mathcal{H}_0 : \mu \geq \mu_0$ in favour of $\mathcal{H}_1 : \mu < \mu_0$ if

$$\bar{X}_n < \mu_0 - \frac{s}{\sqrt{n}} \underbrace{\Phi^{-1}(1-\alpha)}_{\approx 1.64 \text{ for } \alpha = 5\%},$$

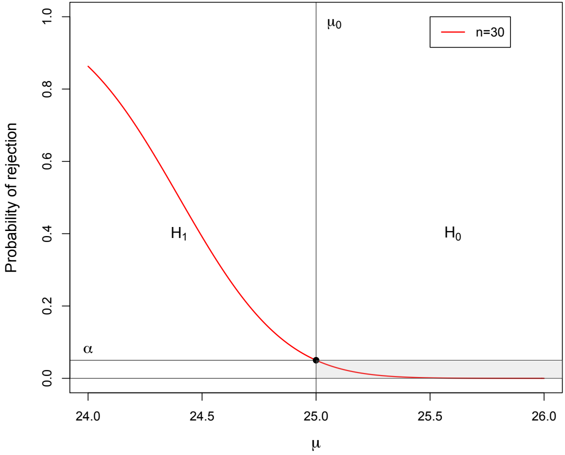
where $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ and $\Phi(z) = P[Z \leq z]$ with $Z \sim \mathcal{N}(0,1)$.



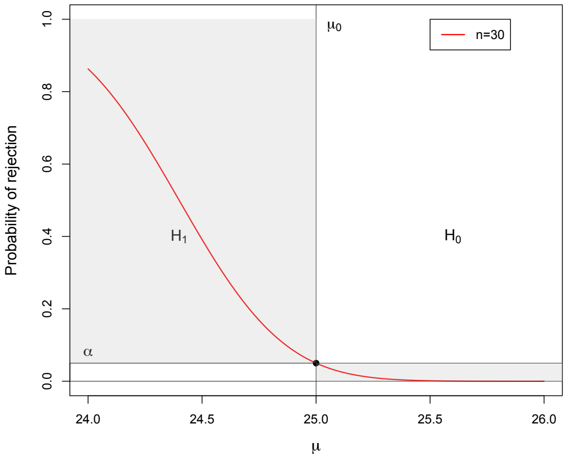
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



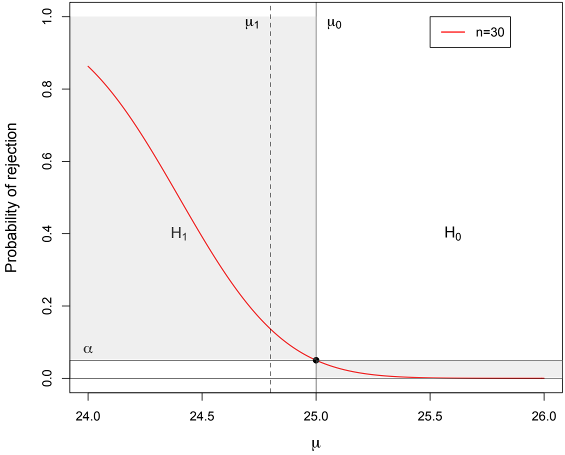
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



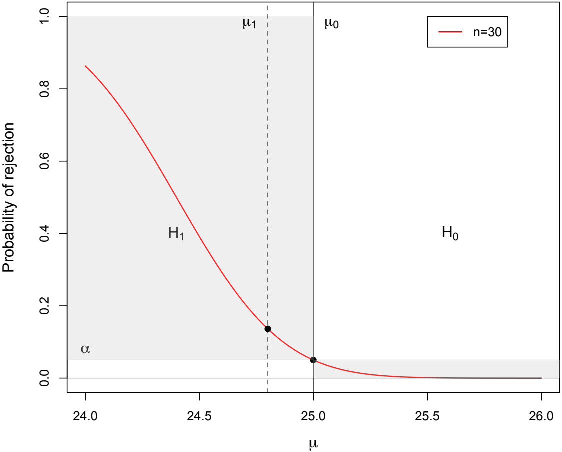
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



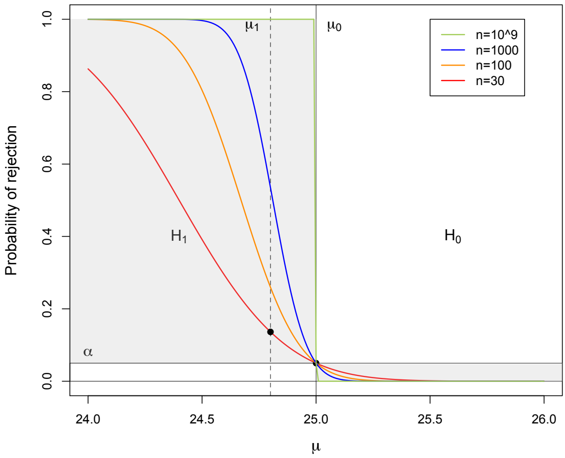
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



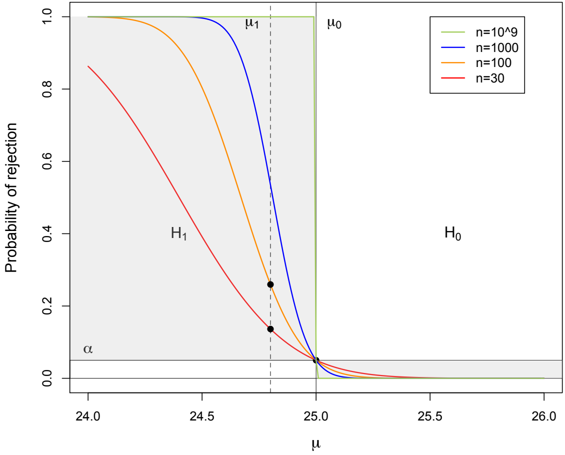
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



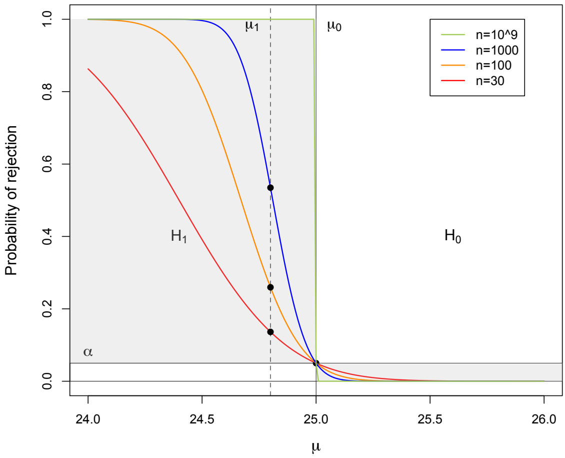
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



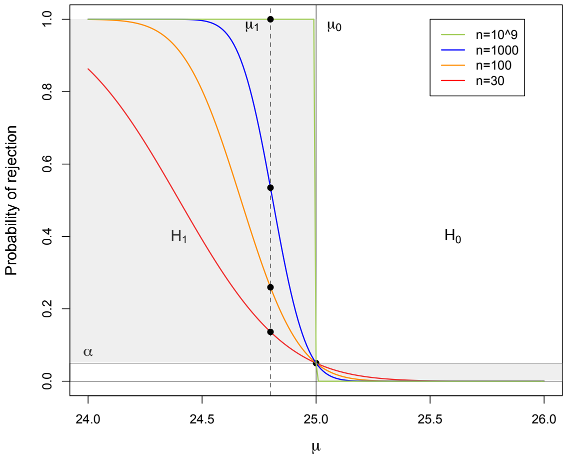
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



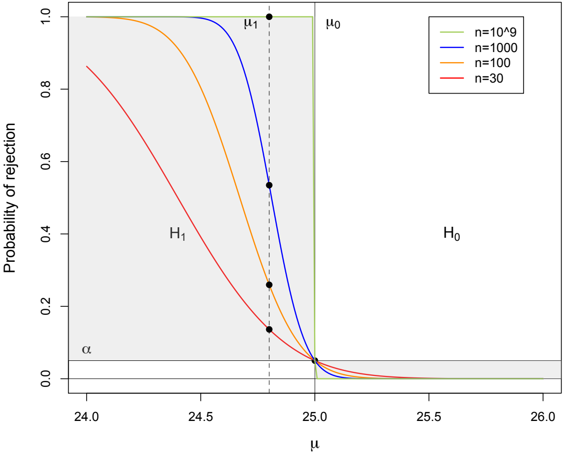
For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



For n large, $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$



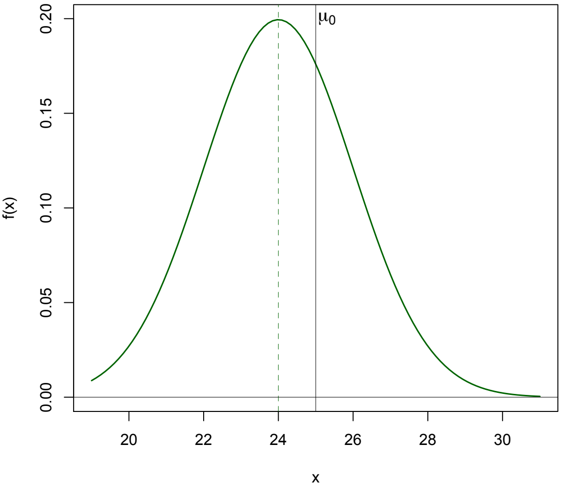
At any $\mu < \mu_0$, then $P_{\mu}[\mathcal{H}_0 \text{ is rejected}] \to 1$ as $n \to \infty$ (consistency)

Consistency holds for all reasonable tests.

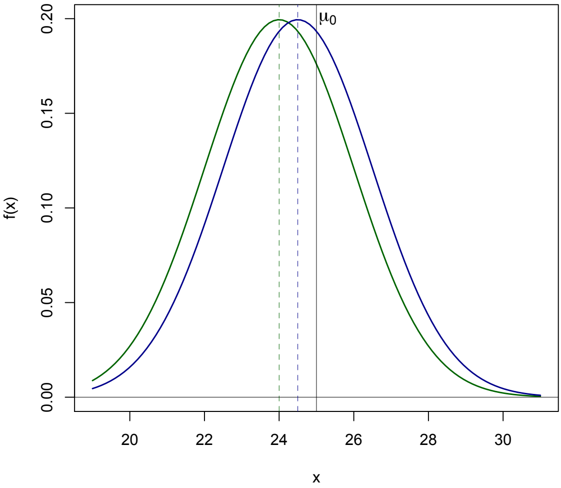
$\rightsquigarrow$ To discriminate between tests, consider increasingly challenging alternatives as the sample size n increases :

$$
\begin{array}{llllll}
X_{11} & & & & & \text{from } f_1 \text{ with mean } \mu_1 = \mu_0 - \nu_1 \tau \\
X_{21} & X_{22} & & & & \text{from } f_2 \text{ with mean } \mu_2 = \mu_0 - \nu_2 \tau \\
\vdots & & \ddots & & & \\
X_{n1} & X_{n2} & \ldots & X_{nn} & & \text{from } f_n \text{ with mean } \mu_n = \mu_0 - \nu_n \tau, \\
\vdots & & & & \ddots &
\end{array}
$$

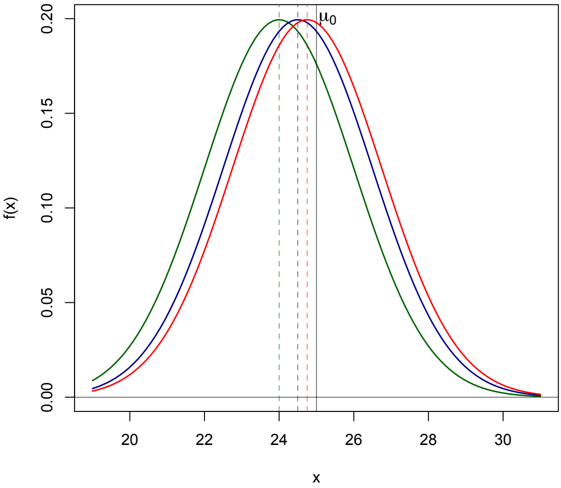
with $(\nu_n) \overset{>}{\to} 0$ and $\tau > 0$.



μ0
0.00
0.05
0.10
0.15
0.20
f(x)
20
22
24
26
28
30
x



0.20
0.15
0.10
0.05
0.00
f(x)
μ0
20
22
24
26
28
30
x



μ_0
0.00
0.05
0.10
0.15
0.20
f(x)
20
22
24
26
28
30
x

Since $P_\mu[\mathcal{H}_0 \text{ is rejected}] \approx \Phi\left(\Phi^{-1}(\alpha) - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma}\right)$ for n large,

$$P_{\mu=\mu_0-\nu_n\tau}[\mathcal{H}_0 \text{ is rejected}] \to \begin{cases} \alpha & \text{if } \nu_n n^{1/2} \to 0 \\ \Phi(\Phi^{-1}(\alpha) + \tau/\sigma) & \text{if } \nu_n = n^{-1/2} \\ 1 & \text{if } \nu_n n^{1/2} \to \infty \end{cases}$$

This test "sees" $n^{-1/2}$-alternatives and is blind to less severe alternatives

This is compatible with consistency (take $\nu_n \equiv 1$)

- Tests may have different consistency rates ________________ (1st order)
- Tests with a common rate may have different accuracies _ (2nd order)

Assume that $f(\mu_0) > 0$. Then, for the test rejecting $\mathcal{H}_0$ if

$$\mathrm{Median}(X_1, \ldots, X_n) < c_\alpha,$$

where c_α is such that the test erroneously rejects $\mathcal{H}_0$ at most with probability α,

$$P_{\mu=\mu_0-\nu_n\tau}[\mathcal{H}_0 \text{ is rejected}] \to \begin{cases} \alpha & \text{if } \nu_n n^{1/2} \to 0 \\ \Phi(\Phi^{-1}(\alpha) + 2f(\mu_0)\tau) & \text{if } \nu_n = n^{-1/2} \\ 1 & \text{if } \nu_n n^{1/2} \to \infty \end{cases}$$

- If $f(\mu_0) > 0$, then this test sees $n^{-1/2}$-alternatives

dominates the $\bar{X}_n$-test iff $2f(\mu_0) > \sigma^{-1}$ (2nd order)

Assume that $f(\mu_0) > 0$. Then, for the test rejecting $\mathcal{H}_0$ if

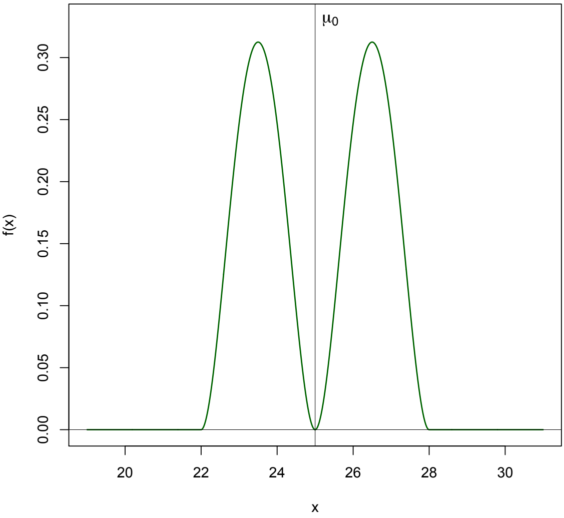
$$\text{Median}(X_1, \ldots, X_n) < c_\alpha,$$

where c_α is such that the test erroneously rejects $\mathcal{H}_0$ at most with probability α,

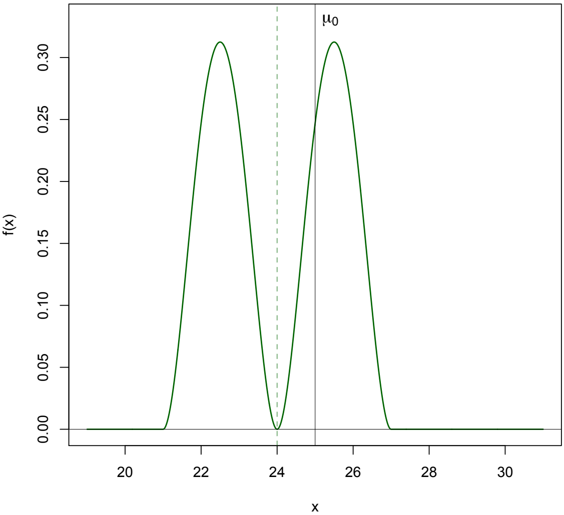
$$P_{\mu=\mu_0-\nu_n\tau}[\mathcal{H}_0 \text{ is rejected}] \to \begin{cases} \alpha & \text{if } \nu_n n^{1/2} \to 0 \\ \Phi(\Phi^{-1}(\alpha) + 2f(\mu_0)\tau) & \text{if } \nu_n = n^{-1/2} \\ 1 & \text{if } \nu_n n^{1/2} \to \infty \end{cases}$$

- If $f(\mu_0) > 0$, then this test sees $n^{-1/2}$-alternatives
  dominates the $\bar{X}_n$-test iff $2f(\mu_0) > \sigma^{-1}$ (2nd order)

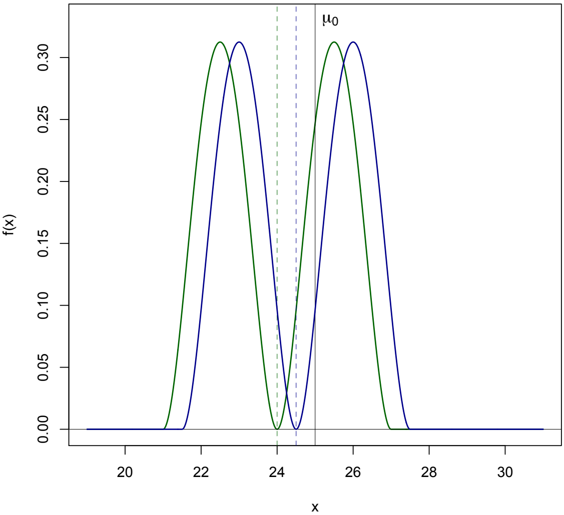
- If $f(\mu_0) = 0$, then this test is blind to $n^{-1/2}$-alternatives! (1st order)
  actually sees $n^{-1/4}$-alternatives only



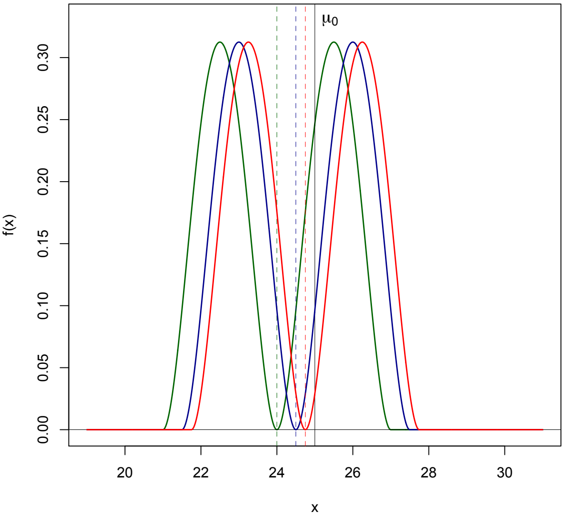
A barman with good days and bad days



A barman with good days and bad days



A barman with good days and bad days



A barman with good days and bad days

For the $\bar{X}_n$-test,

$$P_{\mu=\mu_0-\nu_n\tau}[\mathcal{H}_0 \text{ is rejected}] \to \begin{cases} \alpha & \text{if } \nu_n n^{1/2} \to 0 \\ \Phi(\Phi^{-1}(\alpha)+\tau/\sigma) & \text{if } \nu_n = n^{-1/2} \\ 1 & \text{if } \nu_n n^{1/2} \to \infty \end{cases}$$

This test "sees" $n^{-1/2}$-alternatives and is blind to less severe alternatives

---

Theorem

No tests can see alternatives that are less severe than these alternatives!
(one says that "the $n^{-1/2}$-alternatives are contiguous to $\mathcal{H}_0$")

The $\bar{X}_n$-test is thus rate-optimal

- In standard statistics, one observes

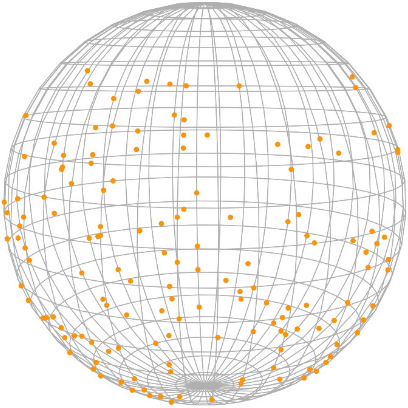
$$X_1, \ldots, X_n$$

with values in $\mathbb{R}^p$. The sample space is the p-dimensional Euclidean space

- In directional statistics, one observes

$$X_1, \ldots, X_n$$

with values in $\mathcal{S}^{p-1} := \{x \in \mathbb{R}^p : \|x\| = \sqrt{x'x} = 1\}$. The sample space is the circle ($p = 2$), the sphere ($p = 3$), or a hypersphere ($p \geq 4$)



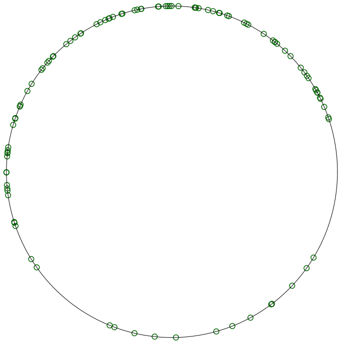
$n = 148$ cosmic ray directions on $\mathcal{S}^2$

Directional statistics find applications in many fields, including

- *Earth sciences* : given the approximately spherical shape of the earth, spherical data arise naturally in earth sciences ($p = 3$)
- *Cosmology* : cosmic ray directions provide spherical data ($p = 3$)
- *Meteorology* : wind directions constitute natural circular data ($p = 2$)
- *Zoology* : study of the moving of animals. In which direction does an ant go when offered a target? ($p = 2$)
- *Linguistics* : clustering texts based on word frequencies ($p >> 3$)
- ...

On the basis of a random sample $X_1, \ldots, X_n$ from a distribution P on $\mathcal{S}^{p-1}$, we want to test

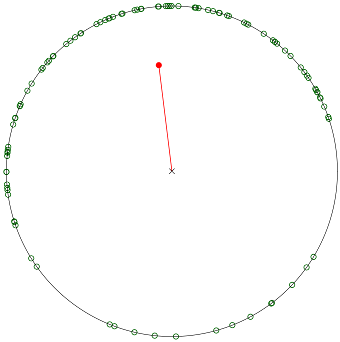
$$\mathcal{H}_0 : P = \mathrm{Unif}(\mathcal{S}^{p-1})$$
$$\mathcal{H}_1 : P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$$



$X_1, \ldots, X_{100}$ from some $P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$

On the basis of a random sample $X_1, \ldots, X_n$ from a distribution P on $\mathcal{S}^{p-1}$, we want to test

$$\mathcal{H}_0 : P = \mathrm{Unif}(\mathcal{S}^{p-1})$$
$$\mathcal{H}_1 : P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$$

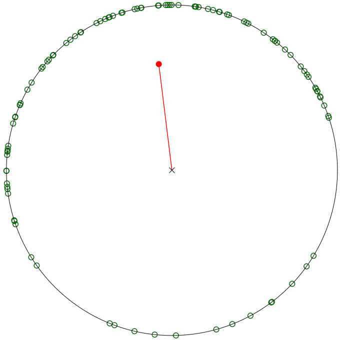


$X_1, \ldots, X_{100}$ from some $P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$

Rayleigh: reject $\mathcal{H}_0$ if $np\|\bar{X}_n\|^2 > \chi^2_{p,1-\alpha}$

On the basis of a random sample $X_1, \ldots, X_n$ from a distribution P on $\mathcal{S}^{p-1}$, we want to test

$$\mathcal{H}_0 : P = \mathrm{Unif}(\mathcal{S}^{p-1})$$
$$\mathcal{H}_1 : P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$$



$X_1, \ldots, X_{100}$ from some $P \neq \mathrm{Unif}(\mathcal{S}^{p-1})$

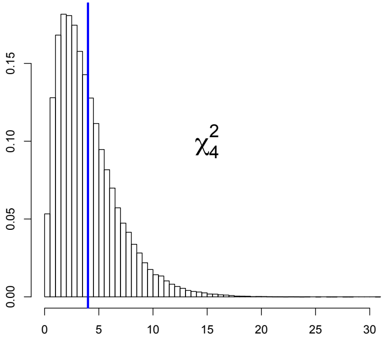
Rayleigh: reject $\mathcal{H}_0$ if $np\|\bar{X}_n\|^2 > \chi^2_{p,1-\alpha}$

Under $\mathcal{H}_0$, $np\|\bar{X}_n\|^2 \xrightarrow{\mathcal{D}} \chi^2_p$ as $n \to \infty$ (implementation requires that $n >> p$)

$Y \sim \chi^2_p$ iff $Y \stackrel{\mathcal{D}}{=} \sum_{i=1}^p Z_i^2$, where the Z_i's are independent $\mathcal{N}(0,1)$

$\mathrm{E}[Y] = p$

$\mathrm{Var}[Y] = 2p$



If $Y_p \sim \chi^2_p$, then the CLT implies that $\dfrac{Y_p - p}{\sqrt{2p}} = \dfrac{Y_p - \mathrm{E}[\chi^2_p]}{\sqrt{\mathrm{Var}[\chi^2_p]}} \xrightarrow[p\to\infty]{} \mathcal{N}(0,1)$

We consider high dimensions (HD), where p is large (and $n >> p$ does not hold)
This requires a double-asymptotic framework: $n \to \infty$ and $p = p_n \to \infty$

$\rightsquigarrow$ We need to consider triangular arrays of the form

$$
\begin{array}{llll l}
X_{11} & & & & \text{with values in } \mathcal{S}^{p_1-1} \\
X_{21} & X_{22} & & & \text{with values in } \mathcal{S}^{p_2-1} \\
\vdots & & \ddots & & \\
X_{n1} & X_{n2} & \ldots & X_{nn} & \text{with values in } \mathcal{S}^{p_n-1} \\
\vdots & & & \ddots &
\end{array}
$$

where, under $\mathcal{H}_0$, observations in row n are sampled from $\mathrm{Unif}(\mathcal{S}^{p_n-1})$

Then $R_n = np_n\|\bar{X}_n\|^2 \xrightarrow[n\to\infty]{\mathcal{D}}$???

Heuristics: the fixed-p asymptotic result

$$R_n \xrightarrow[n\to\infty]{\mathcal{D}} \chi^2_p$$

leads to

$$\frac{R_n - p}{\sqrt{2p}} \xrightarrow[n\to\infty]{\mathcal{D}} \frac{\chi^2_p - p}{\sqrt{2p}} = \frac{\chi^2_p - \mathrm{E}[\chi^2_p]}{\sqrt{\mathrm{Var}[\chi^2_p]}} \xrightarrow[p\to\infty]{} \mathcal{N}(0,1).$$

$\rightsquigarrow$ This suggests that, if $(p_n) \to \infty$ slowly enough,

$$\frac{R_n - p_n}{\sqrt{2p_n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)$$

How fast may p_n go to infinity?

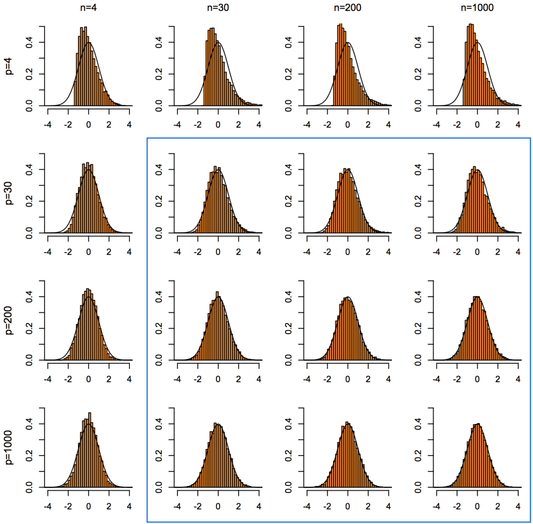
Theorem

If $p_n \to \infty$, then $\dfrac{R_n - p_n}{\sqrt{2p_n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)$ *under* $\mathcal{H}_0$

Theorem

If $p_n \to \infty$, then $$\frac{R_n - p_n}{\sqrt{2p_n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \text{ under } \mathcal{H}_0$$

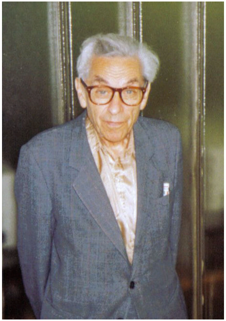
No restrictions on (p_n)!



The proof is based on a CLT for martingale differences; see Billingsley (1995)

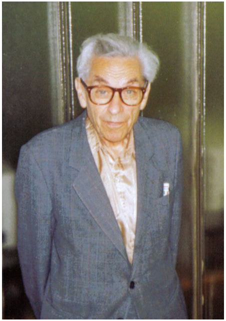
Patrick Paul Billingsley (1925-2011): EB6 (E4-B2)

The proof is based on a CLT for martingale differences; see Billingsley (1995)



Patrick Paul Billingsley (1925-2011): EB6 (E4-B2)

The proof is based on a CLT for martingale differences; see Billingsley (1995)



Patrick Paul Billingsley (1925-2011): EB6 (E4-B2)


AL CAPONE.
He ruled Chicago
with absolute power.
No one could touch him.
No one could stop him.
Until Eliot Ness
and a small force of men
swore they'd bring
him down.
THE
UNTOUCHABLES
PARAMOUNT PICTURES PRESENTS AN ART LINSON PRODUCTION A BRIAN DE PALMA FILM
THE UNTOUCHABLES KEVIN COSTNER CHARLES MARTIN SMITH ANDY GARCIA
ROBERT DE NIRO as AL CAPONE and SEAN CONNERY as MALONE
Music by ENNIO MORRICONE Visual Consultant PATRIZIA VON BRANDENSTEIN Art Director WILLIAM A. ELLIOTT
Director of Photography STEPHEN H. BURUM, A.S.C. Written by DAVID MAMET Produced by ART LINSON Directed by BRIAN DE PALMA
A PARAMOUNT PICTURE

Theorem

If $p_n \to \infty$, then $\dfrac{R_n - p_n}{\sqrt{2p_n}} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,1)$ *under* $\mathcal{H}_0$

The resulting test rejects $\mathcal{H}_0$ if

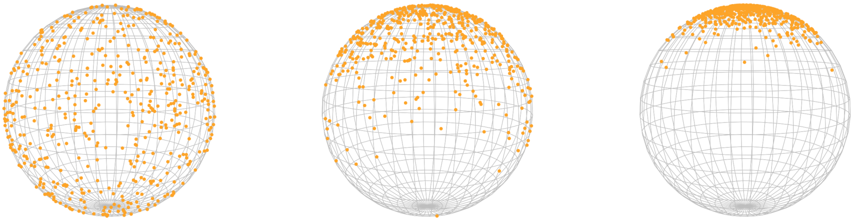
$$\frac{R_n - p_n}{\sqrt{2p_n}} > \Phi^{-1}(1-\alpha)$$

or, equivalently,

$$\frac{R_n - p_n}{\sqrt{2p_n}} > \frac{\chi^2_{p_n,1-\alpha} - p_n}{\sqrt{2p_n}}$$

This is Rayleigh!

Alternatives: FvML densities $x \mapsto c_{p,\kappa} \exp(\kappa\, x'\theta)$, with $\kappa > 0$



θ $\kappa \to 0$ | θ $\kappa = 3$ | θ $\kappa = 10$

The larger κ, the more concentrated the distribution is about θ

$\kappa \to 0$ corresponds to $\mathrm{Unif}(S^{p-1})$

We consider alternatives associated with

$$\begin{array}{llll l} X_{11} & & & & \text{with values in } \mathcal{S}^{p_1-1} \\ X_{21} & X_{22} & & & \text{with values in } \mathcal{S}^{p_2-1} \\ \vdots & & \ddots & & \\ X_{n1} & X_{n2} & \ldots & X_{nn} & \text{with values in } \mathcal{S}^{p_n-1} \\ \vdots & & & \ddots & \end{array}$$

where observations in row n are independent with density

$$x \mapsto c_{p,\kappa_n,f}\, f(\kappa_n\, x'\theta_n),$$

where $\theta_n(\in \mathcal{S}^{p_n-1})$, $\kappa_n(> 0)$, and $f : \mathbb{R} \mapsto \mathbb{R}^+$ monotone $\nearrow$ with $f(0) = f'(0) = 1$

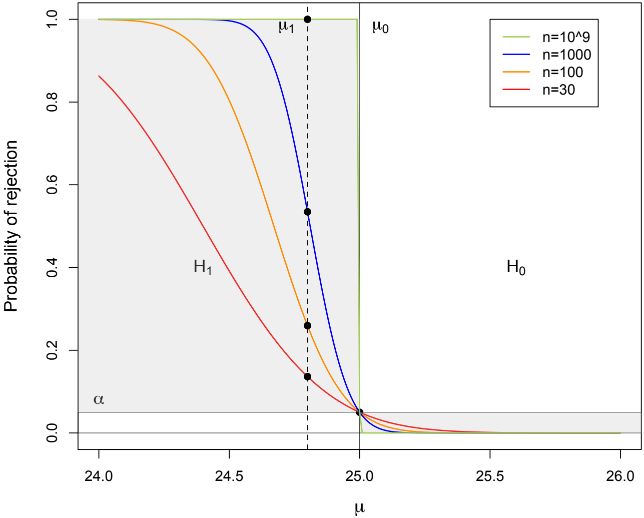
Theorem

Let (p_n) be a sequence of positive integers.
Let (θ_n) be a sequence such that $\theta_n \in \mathcal{S}^{p_n-1}$ for all n.
Let (κ_n) be a positive sequence such that

$$\kappa_n = O\left(\sqrt{\frac{p_n}{n}}\right).$$

Assume that $f''(0)$ exists. Then, these alternatives are contiguous to the null.

- LD: $p_n \equiv p \Rightarrow \kappa_n \sim 1/\sqrt{n}$ provides contiguity
- HD: if $p_n/n \to c > 0 \Rightarrow \kappa_n \equiv \kappa$ provides contiguity
- HHD: if $p_n/n \to \infty \Rightarrow$ some sequences $\kappa_n \to \infty$ provide contiguity



n=10^9
n=1000
n=100
n=30
μ_1
μ_0
H_1
H_0
α
Probability of rejection
0.0
0.2
0.4
0.6
0.8
1.0
24.0
24.5
25.0
25.5
26.0
μ

Theorem

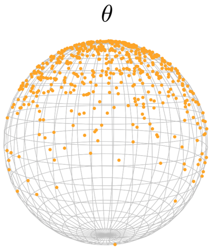
The test rejecting $\mathcal{H}_0$ when

$$\sqrt{np_n}\,\bar{X}_n'\theta_n > \Phi^{-1}(1-\alpha)$$

(i) *satisfies* $P[\mathcal{H}_0 \text{ is rejected}] \to \alpha$ *as* $n \to \infty$ *under* $\mathcal{H}_0$

(ii) *sees the contiguous alternatives with* $\kappa_n = \tau\sqrt{\frac{p_n}{n}}$

(iii) *has maximal asymptotic power among the tests satisfying (i)*



1st-order optimality!

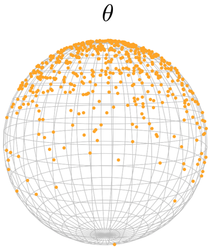
2nd-order optimality!

However, the test rejecting $\mathcal{H}_0$ when

$$\sqrt{np_n}\,\bar{X}_n'\theta_n > \Phi^{-1}(1-\alpha)$$

is an oracle test (addressing the θ-specified problem).

Since θ_n is unspecified in practice, it is natural to replace θ_n with an estimator.



By symmetry, $\bar{X}_n$ is expected to point in the direction of θ

$\rightsquigarrow \hat{\theta}_n = \bar{X}_n/\|\bar{X}_n\|$ is a natural estimator for θ

However, the test rejecting $\mathcal{H}_0$ when

$$\sqrt{np_n}\,\bar{X}_n'\theta_n > \Phi^{-1}(1-\alpha)$$

is an oracle test (addressing the θ-specified problem).

Since θ_n is unspecified in practice, it is natural to replace θ_n with an estimator.

However, the test rejecting $\mathcal{H}_0$ when

$$\sqrt{np_n}\,\bar{X}_n'\theta_n > \Phi^{-1}(1-\alpha)$$

is an oracle test (addressing the θ-specified problem).

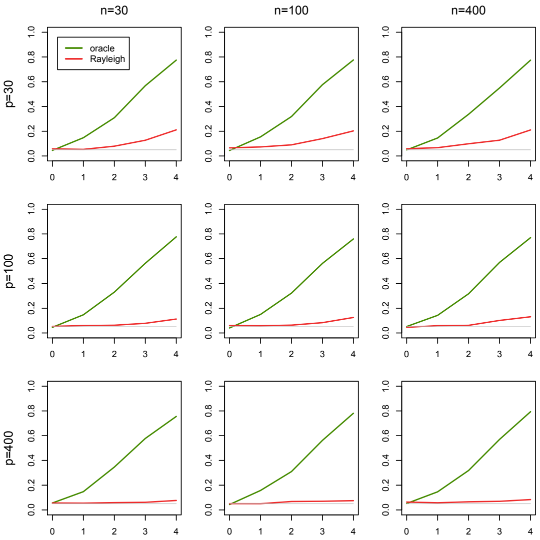
Since θ_n is unspecified in practice, it is natural to replace θ_n with an estimator.

$\rightsquigarrow$ Reject $\mathcal{H}_0$ for large values of

$$\sqrt{np_n}\,\bar{X}_n'\hat{\theta}_n = \sqrt{np_n}\,\bar{X}_n'\big(\bar{X}_n/\|\bar{X}_n\|\big) = \sqrt{np_n}\,\|\bar{X}_n\|,$$

or equivalently, for large values of $np_n\|\bar{X}_n\|^2$ (Rayleigh !!!!!)

Would Rayleigh then be (1st-order and 2nd-order) optimal ?



$\rightsquigarrow$ In HD, Rayleigh is blind to alternatives in $\kappa_n = \tau\sqrt{\frac{p_n}{n}}$

Are there HD alternatives that can be seen by Rayleigh?

Theorem

Let (p_n) be a sequence of positive integers diverging to ∞.
Let (θ_n) be a sequence such that $\theta_n \in \mathcal{S}^{p_n-1}$ for all n.
Then,
(i) if $\kappa_n = \tau p_n^{3/4}/\sqrt{n}$ $(\tau > 0)$, the asymptotic power of Rayleigh is

$$1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{\tau^2}{\sqrt{2}}\right).$$

(ii) if $\kappa_n = o(p_n^{3/4}/\sqrt{n})$, its asymptotic power is α.

In HD, Rayleigh is blind to contiguous alternatives in $\kappa_n = \tau p_n^{1/2}/\sqrt{n}$, but it can see alternatives in $\kappa_n = \tau p_n^{3/4}/\sqrt{n}$.

(In LD, both rates do coincide, so that Rayleigh is rate-optimal)

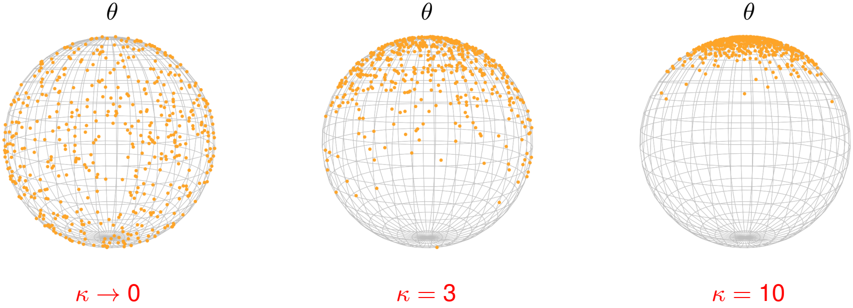
This raises natural questions for the θ-unspecified problem :

- Can a test perform as well as the oracle test ?
- If not, can a test see alternatives in $\kappa_n \sim p_n^{1/2}/\sqrt{n}$?
- If not, can a test see alternatives that are less severe than $\kappa_n \sim p_n^{3/4}/\sqrt{n}$?
- If not, would Rayleigh be not only 1st-order optimal but also 2nd-order optimal ?

If one restricts to rotation-invariant tests, we can answer the questions :

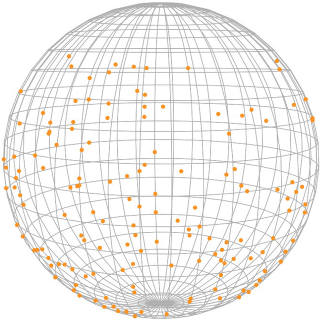
- Can a test perform as well as the oracle test ?
  No
- If not, can a test see alternatives in $\kappa_n \sim p_n^{1/2}/\sqrt{n}$?
  No
- If not, can a test see alternatives that are less severe than $\kappa_n \sim p_n^{3/4}/\sqrt{n}$?
  No
- If not, would Rayleigh be not only 1st-order optimal but also 2nd-order optimal ?
  Yes

$X_1, \ldots, X_n$ randomly sampled from a density $x \mapsto c_{p,\kappa} f(\kappa\, x'\theta)$ over $\mathcal{S}^{p-1}$,
where $\theta \in \mathcal{S}^{p-1}$ (signal direction), $\kappa > 0$ (signal strength)
$f : \mathbb{R} \mapsto \mathbb{R}^+$ monotone $\nearrow$, with $f(0) = f'(0) = 1$



For κ small, inference on θ is challenging

We consider inference on θ (i.e., on the mode/direction of the signal) in the neighbourhood of $\kappa = 0$ (i.e., when the signal is weak)



$n = 148$ cosmic ray directions on $\mathcal{S}^2$

κ small

We consider triangular arrays of the form

$$\begin{array}{llll l}
X_{11} & & & & \text{with values in } \mathcal{S}^{p-1} \\
X_{21} & X_{22} & & & \text{with values in } \mathcal{S}^{p-1} \\
\vdots & & \ddots & & \\
X_{n1} & X_{n2} & \ldots & X_{nn} & \text{with values in } \mathcal{S}^{p-1} \\
\vdots & & & & \ddots
\end{array}$$

where observations in row n are independent with density $x \mapsto c_{p,\kappa_n,f}\, f(\kappa_n\, x'\theta_n)$, with $\theta_n(\in \mathcal{S}^{p-1})$, $\kappa_n \to 0$, and $f : \mathbb{R} \mapsto \mathbb{R}^+$ monotone $\nearrow$ with $f(0) = f'(0) = 1$

A point estimation result for $\hat{\theta}_n = \bar{X}_n / \|\bar{X}_n\|$...

Theorem

Under $\mathrm{P}^{(n)}_{\theta,\kappa_n,f}$, *with* $\kappa_n = \sqrt{p}\xi\eta_n$, *we have the following:*

(i) For $\eta_n \equiv 1$,

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0\,,\, \frac{C_{p,f}}{\xi^2}(I_p - \theta\theta')\right) \quad \textit{for some } C_{p,f} > 0;$$

(ii) For $\eta_n \to 0$ *with* $\sqrt{n}\,\eta_n \to \infty$,

$$\sqrt{n}\,\eta_n(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0\,,\, \frac{1}{\xi^2}(I_p - \theta\theta')\right)$$

(iii) For $\sqrt{n}\,\eta_n \to 1$,

$$\hat{\theta}_n \xrightarrow{\mathcal{D}} \frac{Z}{\|Z\|}, \quad \textit{with } \; Z \sim \mathcal{N}\big(\xi\theta, I_p\big)$$

(iv) For $\sqrt{n}\,\eta_n \to 0$, $\hat{\theta}_n \xrightarrow{\mathcal{D}}$ *the uniform distribution over* $\mathcal{S}^{p-1}$.

We want to test $\mathcal{H}_0 : \theta = \theta_0$, where θ_0 is fixed.

- The Wald test rejects $\mathcal{H}_0$ if

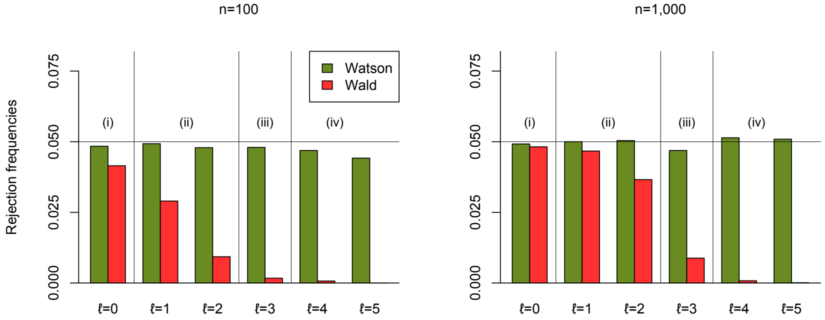
$$S_n := \frac{n(p-1)(\bar{X}_n'\theta_0)^2}{1 - \frac{1}{n}\sum_{i=1}^n (X_i'\theta_0)^2} (\hat{\theta}_n - \theta_0)'(I_p - \theta_0\theta_0')(\hat{\theta}_n - \theta_0) > \chi^2_{p-1,1-\alpha},$$

  with $\hat{\theta}_n = \bar{X}_n / \|\bar{X}_n\|$

- The Watson rejects $\mathcal{H}_0$ if

$$W_n := \frac{n(p-1)}{1 - \frac{1}{n}\sum_{i=1}^n (X_i'\theta_0)^2} \|(I_p - \theta_0\theta_0')\bar{X}_n\|^2 > \chi^2_{p-1,1-\alpha}.$$

For κ fixed, both tests are asymptotically equivalent



Null rejection frequencies of Watson and Wald (at $\alpha = 5\%$)

FvML case with $\kappa = \kappa_\ell = \sqrt{p}n^{-\ell/6}$

Theorem

With $\kappa_n = \sqrt{p}\xi\eta_n + o(\eta_n)$, we have the following as $n \to \infty$ under $\mathrm{P}^{(n)}_{\theta_0,\kappa_n,f}$:

(i) if $\eta_n \equiv 1$, then

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}$$

(ii) if $\eta_n \to 0$ with $\sqrt{n}\,\eta_n \to \infty$, then

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}$$

(iii) if $\sqrt{n}\,\eta_n \to 1$, then

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1 + \frac{Q}{(Z+\xi)^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

(iv) if $\sqrt{n}\,\eta_n \to 0$, then

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1 + \frac{Q}{Z^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

Theorem

With $\kappa_n = \sqrt{p}\xi\eta_n + o(\eta_n)$, *we have the following as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\theta_0,\kappa_n,f}$ *:*

(i) if $\eta_n \equiv 1$, *then*

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}$$

(ii) if $\eta_n \to 0$ *with* $\sqrt{n}\,\eta_n \to \infty$, *then*

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}$$

(iii) if $\sqrt{n}\,\eta_n \to 1$, *then*

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1 + \frac{Q}{(Z+\xi)^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

(iv) if $\sqrt{n}\,\eta_n \to 0$, *then*

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1 + \frac{Q}{Z^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

Watson is robust to arbitrary small departures from uniformity

Wald is not

Of course such robustness should not be obtained at the expense of power...

Theorem

With $\kappa_n = \sqrt{p}\xi\eta_n + o(\eta_n)$ *and* $\tau_n \to \tau$, *we have that, as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\theta_0+\nu_n\tau_n,\kappa_n,f}$,

(i) if $\eta_n \equiv 1$, *then, for* $\nu_n = 1/\sqrt{n}$,

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{\xi^2}{C_{p,f}}\|\tau\|^2\right)$$

(ii) if $\eta_n \to 0$ *with* $\sqrt{n}\,\eta_n \to \infty$, *then, for* $\nu_n = 1/(\sqrt{n}\,\eta_n)$,

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}(\xi^2\|\tau\|^2)$$

(iii) if $\sqrt{n}\,\eta_n \to 1$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{1}{4}\xi^2\|\tau\|^2(4-\|\tau\|^2)\right) \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \ldots$$

(iv) if $\sqrt{n}\,\eta_n \to 0$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1+\frac{Q}{Z^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

Theorem

With $\kappa_n = \sqrt{p}\xi\eta_n + o(\eta_n)$ *and* $\tau_n \to \tau$, *we have that, as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\theta_0+\nu_n\tau_n,\kappa_n,f}$,

(i) if $\eta_n \equiv 1$, *then, for* $\nu_n = 1/\sqrt{n}$,

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{\xi^2}{C_{p,f}}\|\tau\|^2\right)$$

(ii) if $\eta_n \to 0$ *with* $\sqrt{n}\,\eta_n \to \infty$, *then, for* $\nu_n = 1/(\sqrt{n}\,\eta_n)$,

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}(\xi^2\|\tau\|^2)$$

(iii) if $\sqrt{n}\,\eta_n \to 1$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{1}{4}\xi^2\|\tau\|^2(4-\|\tau\|^2)\right) \quad \text{and} \quad S_n \xrightarrow{\mathcal{D}} \ldots$$

(iv) if $\sqrt{n}\,\eta_n \to 0$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \text{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1+\frac{Q}{Z^2}\right)^{-1} Q \quad \text{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$

Theorem

With $\kappa_n = \sqrt{p}\xi\eta_n + o(\eta_n)$ *and* $\tau_n \to \tau$, *we have that, as* $n \to \infty$ *under* $\mathrm{P}^{(n)}_{\theta_0+\nu_n\tau_n,\kappa_n,f}$,

(i) if $\eta_n \equiv 1$, *then, for* $\nu_n = 1/\sqrt{n}$,

$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{\xi^2}{C_{p,f}}\|\tau\|^2\right)$$

(ii) if $\eta_n \to 0$ *with* $\sqrt{n}\,\eta_n \to \infty$, *then, for* $\nu_n = 1/(\sqrt{n}\,\eta_n)$,

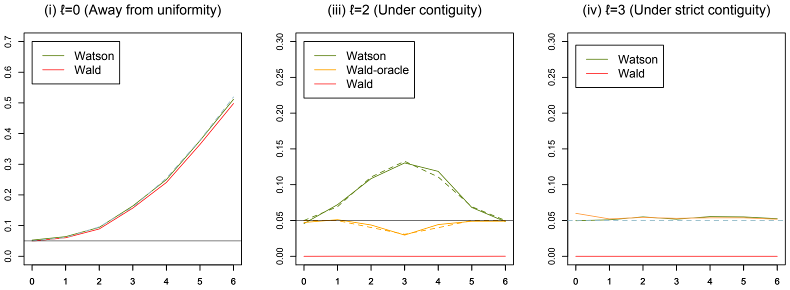
$$W_n = S_n + o_{\mathrm{P}}(1) \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\xi^2\|\tau\|^2\right)$$

(iii) if $\sqrt{n}\,\eta_n \to 1$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1}\left(\frac{1}{4}\xi^2\|\tau\|^2(4-\|\tau\|^2)\right) \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \ldots$$

(iv) if $\sqrt{n}\,\eta_n \to 0$, *then, for* $\nu_n \equiv 1$,

$$W_n \xrightarrow{\mathcal{D}} \chi^2_{p-1} \quad \textit{and} \quad S_n \xrightarrow{\mathcal{D}} \left(1+\frac{Q}{Z^2}\right)^{-1} Q \quad \textit{with} \quad \begin{array}{l} Q \sim \chi^2_{p-1} \\ Z \sim \mathcal{N}(0,1) \end{array} \quad \perp\!\!\!\perp$$



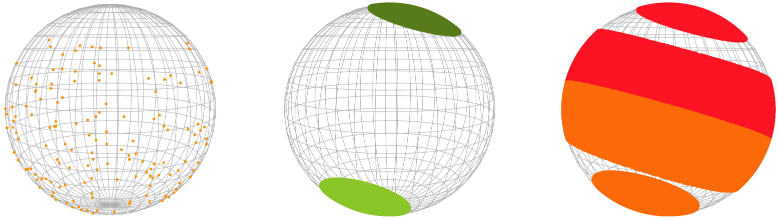
Non-null rejection frequencies of Watson and Wald (at $\alpha = 5\%$)
FvML case with $\kappa = \kappa_\ell = \sqrt{p}n^{-\ell/6}$

This raises natural questions, in any given regime (i)-(iv):

- Is there a test that can see less severe alternatives than Watson can?
  No! Watson is adaptively 1st-order optimal

- If not, is Watson even 2nd-order optimal?
  Yes! But in a sense that depends on the regime :

  In (i), Watson is 2nd-order optimal for $f = \exp$ only
  In (ii), Watson is 2nd-order optimal uniformly in f
  In (iii), Watson is 2nd-order optimal uniformly in f but only locally-in-τ
  In (iv), Watson is 2nd-order optimal uniformly in f, but in a degenerate way



Watson and Wald 95% confidence zones for θ

Cutting, C., Paindaveine, D., and Verdebout, T. (2017). Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. *Annals of Statistics*, to appear.

Paindaveine, D., and Verdebout, T. (2017). Inference on the mode of weak directional signals : A Le Cam perspective on hypothesis testing near singularities. *Annals of Statistics* 45, 800-832.

Thank you for your attention