Mathematical Statistics II

STA2212H S LEC0101

Week 9

March 14 2023

HEALTH

A New Turn in the Fight Over Masks

A crucial pandemic question is deceptively hard to answer.

By Yasmin Tayag

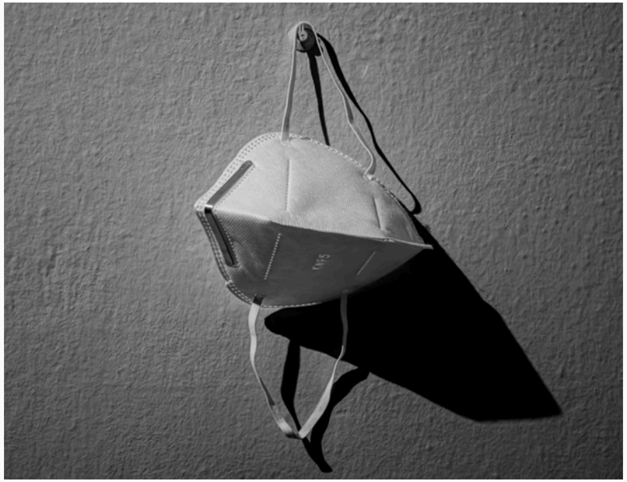
Masks

HEALTH

A New Turn in the Fight Over Masks

A crucial pandemic question is deceptively hard to answer.

By Yasmin Tayag



OPINION

ZEYNEP TUFEKCI

Here's Why the Science Is Clear That Masks Work

March 10, 2023 8 MIN READ

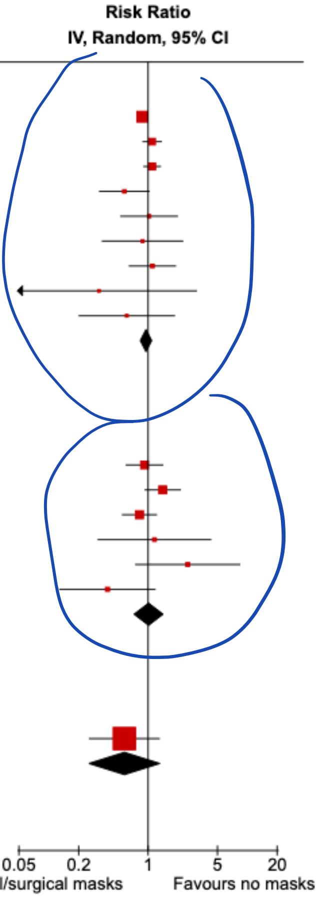
Atlantic (Y. Tayag)

NY Times (Z. Tufekci)

Masks

Cochrane Review

Study or Subgroup	log[RR]	SE	Medical/surgical masks Total	No masks Total	Weight	Risk Ratio IV, Random, 95% CI
1.1.1 Influenza/COVID-like illness						
Abaluck 2022 (1)	-0.135	0.036	111525	155268	41.4%	0.87 [0.81 , 0.94]
Aiello 2012	0.095	0.115	392	370	19.8%	1.10 [0.88 , 1.38]
Alfelali 2020	0.095	0.105	3864	3823	21.9%	1.10 [0.90 , 1.35]
Barasheed 2014	-0.55	0.3	75	89	4.6%	0.58 [0.32 , 1.04]
Canini 2010	0.025	0.342	148	158	3.6%	1.03 [0.52 , 2.00]
Cowling 2008	-0.128	0.483	61	205	1.9%	0.88 [0.34 , 2.27]
MacIntyre 2009	0.1	0.28	186	100	5.2%	1.11 [0.64 , 1.91]
MacIntyre 2016	-1.139	1.16	302	295	0.3%	0.32 [0.03 , 3.11]
Suess 2012	-0.494	0.571	26	30	1.4%	0.61 [0.20 , 1.87]
Subtotal (95% CI)			**116579**	**160338**	**100.0%**	**0.95 [0.84 , 1.09]**
Heterogeneity: Tau² = 0.01; Chi² = 11.44, df = 8 (P = 0.18); I² = 30%						
Test for overall effect: Z = 0.71 (P = 0.48)						
1.1.2 Laboratory-confirmed influenza or SARS-cov-2						
Aiello 2012	-0.083	0.223	392	370	25.9%	0.92 [0.59 , 1.42]
Alfelali 2020	0.34	0.215	3864	3823	26.7%	1.40 [0.92 , 2.14]
Bundgaard 2021 (2)	-0.2	0.208	2392	2470	27.4%	0.82 [0.54 , 1.23]
Cowling 2008	0.148	0.674	61	205	5.8%	1.16 [0.31 , 4.34]
MacIntyre 2009	0.92	0.6225	186	100	6.6%	2.51 [0.74 , 8.50]
Suess 2012	-0.942	0.57	26	30	7.7%	0.39 [0.13 , 1.19]
Subtotal (95% CI)			**6921**	**6998**	**100.0%**	**1.01 [0.72 , 1.42]**
Heterogeneity: Tau² = 0.07; Chi² = 8.52, df = 5 (P = 0.13); I² = 41%						
Test for overall effect: Z = 0.07 (P = 0.95)						
1.1.3 Laboratory-confirmed other respiratory viruses						
Bundgaard 2021	-0.55	0.42	2392	2470	100.0%	0.58 [0.25 , 1.31]
Subtotal (95% CI)			**2392**	**2470**	**100.0%**	**0.58 [0.25 , 1.31]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.31 (P = 0.19)						



Footnotes

(1) Covid-like-illness

(2) SARS-cov-2

Today

1. Next lectures
2. Recap
3. Robust estimation
4. Asymptotic theory

Upcoming

- March 20 3.30 – 4.30 DoSS 9014 & online Details
  "Using Data Science to Optimize Business – Opportunities & Challenges"

  Alison Burnham, Digitization Office, RepairSmith Inc

Next Lectures

- March 14 10.00 – 13.00
- March 21 11.00 – 13.00
- March 28 10.00 – 12.00
- April 4 Project presentations

Next Lectures

- March 14 10.00 – 13.00
- March 21 11.00 – 13.00
- March 28 10.00 – 12.00
- April 4 Project presentations

Week	Date	Methods	References
1	Jan 10	Likelihood inference: review of ML estimation; mis-specified models; computation; nonparametric mle	MS §§5.1–7, SM Ch 4
2	Jan 17	Bayesian estimation; Bayesian inference	MS §5.8; AoS §§ 11.1–4; SM §§11.1,2
3	Jan 24	Optimality in estimation	MS Ch 6; AoS Ch 12; SM §7.1, 11.5.2
4	Jan 31	Interval estimation; Confidence bands	MS §§7.1,2; AoS Ch 7; SM §7.1.4
5	Feb 7	Hypothesis testing; likelihood ratio tests	MS §§7.1–4 AoS Ch 10.6, SM
6	Feb 14	Significance testing	MS §7.5; AoS §10.2,6; SM Ch 4, §7.3.1
	Feb 21	Break	
7	Feb 28	Significance testing	SM 7.3.1
7	Feb 28	Goodness-of-fit testing	MS Ch 9; AoS §§10.3,4,5,8; SM p.327-8 (hard)
8	Mar 7	Multiple testing and FDR	AoS Ch 10.7, EH Ch 15.1,2
9	Mar 14	Robust Estimation Likelihood Asymptotics	MS 8.4, 8.6; SM 8.4 SM 4.4, 4.5
10	Mar 21	Causal Inference	AoS 16, 17
11	Mar 28	Classification	AoS 22
12	Apr 4	Course Summary; Presentations	

References

MS: *Mathematical Statistics* by K. Knight (Chapman & Hall/CRC).

AoS: *All of Statistics* by L. Wasserman (Springer) If your copy has a **Chapter 1. Introduction**, then all Chapter numbers increase by 1.

SM: *Statistical Models* by A.C. Davison (Cambridge University Press)

Recap

- multiple testing: family-wise error rate (FWER); false discovery rate (FDR); Benjamini-Hochberg method controls FDR

- goodness-of-fit tests based on empirical cdf $\hat{F}_n(\cdot)$
  - Kolmogorov-Smirnov
  - Cramer-vonMises
  - Anderson-Darling
- Brownian bridge; limit distributions

- goodness-of-fit tests based on multinomial distribution

Benjamini-Hochberg

smallest ↑ largest (HW8) $t_{(i)}$ largest ↓ smallest

- order the p-values $p_{(1)}, \ldots, p_{(m)}$
- find i_{max}, the largest index for which

$$p_{(i)} \leq \frac{i}{m} q$$

- Let BH_q be the rule that rejects H_{0i} for $i \leq i_{max}$, not rejecting otherwise

Benjamini-Hochberg

- order the p-values $p_{(1)}, \ldots, p_{(m)}$
- find i_{max}, the largest index for which

$$p_{(i)} \leq \frac{i}{m} q$$

- Let BH_q be the rule that rejects H_{0i} for $i \leq i_{max}$, not rejecting otherwise
- change the bound under dependence

$$p_{(i)} \leq \frac{i}{mC_m} q \qquad C_m = \sum_{i=1}^{m} \frac{1}{i}$$

Benjamini-Hochberg

- order the p-values $p_{(1)}, \ldots, p_{(m)}$
- find i_{max}, the largest index for which

$$p_{(i)} \leq \frac{i}{m} q$$

- Let BH_q be the rule that rejects H_{0i} for $i \leq i_{max}$, not rejecting otherwise
- change the bound under dependence

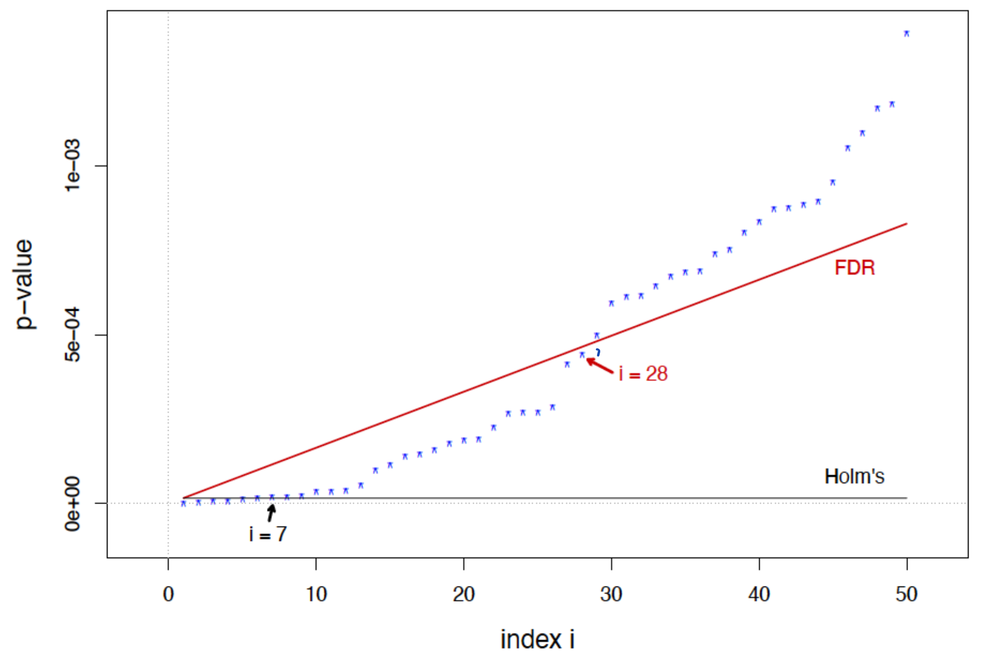
$$p_{(i)} \leq \frac{i}{mC_m} q \qquad C_m = \sum_{i=1}^{m} \frac{1}{i}$$

- Theorem: If the p-values corresponding to valid null hypotheses are independent of each other, then

$$FDR(BH_q) = \pi_0 q \leq q, \qquad \text{where } \pi_0 = \underline{m_0/m} \qquad 0 < q < 1$$

prob(null)

π_0 unknown but close to 1



$q = 0.1$

BH-q: reject H_{0i} for $p_{(i)} \leq \frac{i}{m} q$

$p_{(1)} \leq \cdots \leq p_{(m)}$

BH-q: reject H_{0i} for $p_{(i)} \le \frac{i}{m} q$ $\quad$ $FDR(BH_q) = \pi_0 q \le q$, $\quad$ where $\pi_0 = m_0/m$

$0 < t \le 1$:

$$
\begin{aligned}
R(t) &= \#\{p_i \le t\} \\
V(t) &= \#\{p_i \le t, H_{0i} \text{ true }\} \\
FDP(t) &= V(t)/\max(R(t), 1) \\
Q(t) &= mt/\max(R(t), 1) \\
t_q &= \sup_t \{Q(t) \le q\}
\end{aligned}
$$

		H_0 not rejected	H_0 rejected	
truth	H_0 true	U	V	m_0
	H_1 true	T	S	m_1
		$m - R$	R	m

1. $R(p_{(i)}) = i = Q(p_{(i)}) = \frac{m p_{(i)}}{i} \le q \Rightarrow p_{(i)} \le t_q$
2. $A(t) = \frac{V(t)}{t}$ $\quad$ $E\{A(s) \mid A(t)\} = A(t)$, $s \le t$ $\Rightarrow A(\cdot)$ m'gale

   $E\{A(t_q)\} = EA(1) = \frac{EV(1)}{1} = m_0$
3.

Proof

BH-q: reject H_{0i} for $p_{(i)} \le \frac{i}{m} q$

$FDR(BH_q) = \pi_0 q \le q$, where $\pi_0 = m_0/m$

$0 < t \le 1$:

$$
\begin{aligned}
R(t) &= \#\{p_i \le t\} \\
V(t) &= \#\{p_i \le t, H_{0i} \text{ true }\} \\
FDP(t) &= V(t)/\max(R(t), 1) \\
\boxed{Q(t) = mt/\max(R(t), 1)} \\
t_q &= \sup_t \{Q(t) \le q\}
\end{aligned}
$$

$E\left\{\frac{V}{\max(R,1)}\right\}$

		H_0 not rejected	H_0 rejected	
	H_0 true	U	V	m_0
truth				
	H_1 true	T	S	m_1
		$m - R$	R	m

$R(p_{(i)}) = i \implies Q(p_{(i)}) = mp_{(i)}/i$

1. BH-q $\iff$: reject H_{0i} for $\boxed{p_{(i)} \le t_q}$
2. $A(t) = V(t)/t$, $\mathrm{E}\{A(s) \mid A(t)\} = A(t), s \le t$
3. $\max\{R(t_q), 1\} = \frac{mt_q}{Q(t_q)} = \frac{mt_q}{q} \implies FDP(t_q) = \frac{q}{m}\frac{V(t_q)}{t_q} = q\frac{m_0}{m}$

if $p_{(i)} \le \frac{i}{m} q$ then $p_{(i)} \le t_q$ $= EV(t_q)/t_q$

$\implies \boxed{\mathrm{E}\{A(t_q)\} = E\{A(1)\} = m_0}$

$EV(t_q) = m_0$

2. $E\{A(s) \mid A(t)\} = A(t), \quad s \leq t$

$A(t) = V(t)/t = \#\{p_i \leq t, H_{0i} \text{ true }\}/t$

$\cancel{X_1, \ldots, X_n \text{ i.i.d.}} \sim U(0,1)$

$s < t$

$\text{pr}(X \leq s \mid X \leq t) = \cdots = \dfrac{s}{t}$

$= \dfrac{P_1(X \leq s,\ X \leq t)}{P(X \leq t)} = \dfrac{P_1(X \leq s)}{P(X \leq t)}$

$= \dfrac{s}{t} \quad 0 < s < t$

$U(0, \frac{s}{t})$

use this?

$X_1, \ldots X_m$ iid $U(0,1)$ to go $\rightarrow$ $E(\quad)$

- Linear regression $Y = X\beta + \sigma\epsilon$ $E(\epsilon_n) = 0, \text{var}(\epsilon) = I_n$

$_{n\times 1}$ $X_{n\times p}$ $\beta_{p\times 1}$

$E(Y|X) = X\beta$

$\text{cov}(Y|X) = \sigma^2 I_n$

- Least squares estimator

$$\hat{\beta}_{LS} = \underbrace{(X^TX)^{-1}X^T}_{A}Y$$

(X full rank)

- Gauss-Markov theorem: BLUE; MVUE

$$\text{var}(\hat{\beta}_{LS}) \leq \text{var}(\tilde{\beta}),$$

$\tilde{\beta} = AY$ for some $p\times n$ matrix (non-random) A

- fun fact: if $\epsilon_i \sim t_\nu$ are independent,

$\nu > 2$

$$\text{var}(\hat{\beta}_{LS}) = \sigma^2(X^TX)^{-1}\frac{\nu}{\nu-2}; \quad \text{a.var}(\hat{\beta}_{MLE}) = \sigma^2(X^TX)^{-1}\frac{\nu+3}{\nu+1}$$

(not AY)

- $\hat{\beta}_{LS}$ has asymptotic relative efficiency $\dfrac{(\nu-2)(\nu+3)}{\nu(\nu+1)}$

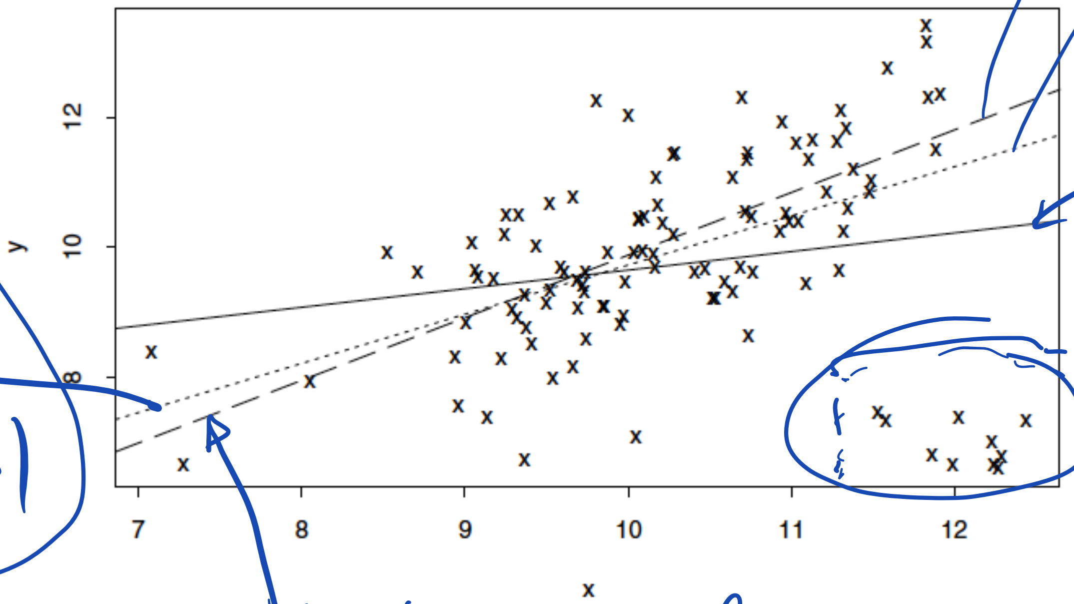
$\nu = 5, 10, 20$; $\textit{eff} = .80, .95, .99$

heavy tails

$\approx \frac{1}{2}$ residuals above / below line

resistant / robust to contamination

$\min_{\beta} |y - X\beta| = \min_{\beta} \sum |y_i - x_i^T \beta|$



LS line $\hat{y}_i = \hat{\beta}_1 x_i$

susceptible to outliers

least median of squares $\min_{\beta} \operatorname{med}\{(y_i - x_i^T \beta)^2\}$

Figure 8.2 *Estimated regression lines; the solid line is the least squares line, the dotted line is the L_1 line and the dashed line is the LMS line. Notice how the least squares line is pulled more towards the 10 "outlying" observations than are the other two lines.*

$\min_{\mu} \sum_{i=1}^{n} |y_i - \mu|$

$\hat{\mu} = \operatorname{med}(y_i)$

loss f=

- Linear regression: $y = X\beta + \sigma\epsilon$

$$\min_{\beta} \sum_{i=1}^{n} \underline{(y_i - x_i^T\beta)^2} \longrightarrow \min_{\beta} \sum_{i=1}^{n} \underline{\rho(y_i - x_i^T\beta)}$$

$\rho(\cdot)$ to be chosen

- or more typically, $$\min_{\beta} \sum_{i=1}^{n} \rho\{(y_i - x_i^T\beta)/\sigma\}$$

← t-dist f=

- various choices for $\rho(u)$: $u^2/2$, $|u|$, $\nu \log(1 + u^2/\nu)/2$,

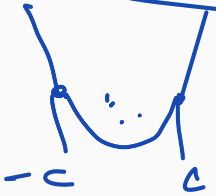
tuning

$$\text{Huber: } \rho(u) = \begin{cases} u^2, & |u| \le c \\ c(2|u| - c) & \text{otherwise} \end{cases}$$

- MS also considers Least Median Squares estimator

robustness

- Tukey's biweight



effic

smooth

geom → smooth

SM $\rho'(u)$ get 0 weight in loss f^n

$Y = X\beta + \sigma\varepsilon$

But not assuming dist^n for ε

$$\min_{\beta} \sum_{i=1}^{n} \rho\{(y_i - x_i^T\beta)/\sigma\}$$

$\rho(\cdot)$ convex f^n

$\hat{\beta}$ sol^n

- Theorem: if $\rho(\cdot)$ is convex, and MS Thm 8.6

$$g = \psi(t) = \rho'(t),$$

$E[-\psi'(\varepsilon)]^{-1} E\{\psi(\varepsilon)\cdot\psi(\varepsilon)^T\} E[-\psi'(\varepsilon)]^{-1}$

E cov E

is non-decreasing, then

$$\max_{1\le i\le n} x_i^T(X^TX)^{-1}x_i \to 0, \quad n\to\infty$$

$$\implies A_n(\hat{\beta}-\beta) \xrightarrow{d} N_p(0, \gamma^2 I), \qquad \gamma^2 = \frac{\mathrm{E}\{\psi^2(\epsilon)\}}{\mathrm{E}^2\{\psi'(\epsilon)\}}$$

← var

← E^2

$A_n^2 = (X_n^T X_n)$

[simple linearity

Aside: ordinary least squares

$$\hat{\beta}_{n,LS} = (X_n^T X_n)^{-1} X_n^T Y_n$$

← want asympt result on $\hat{\beta}$

$$\hat{\beta}_{n,LS} - \beta = (X_n^T X_n)^{-1} X_n^T \epsilon_n$$

$$\boxed{A_n^2 = X_n^T X_n}$$

$$A_n(\hat{\beta}_{n,LS} - \beta) = A_n^{-1} X_n^T \epsilon_n$$

p fixed ~~as~~ $n \to \infty$

$$\max_{1 \le i \le n} x_i^T (X_n^T X_n)^{-1} x_i \to 0 \quad n \to \infty \quad \Longrightarrow \quad A_n(\hat{\beta}_{n,LS} - \beta) \xrightarrow{d} N(0, \sigma^2 I)$$

$\hat{\beta}_{n,LS} \stackrel{\sim}{\cdot} N(\beta, \sigma^2 (X^T X)^{-1})$

Note that $x_i^T (X_n^T X_n)^{-1} x_i = h_{ii}$, where $H_n = X_n (X_n^T X_n)^{-1} X_n^T$ and trace$(H) = p$

Simple linear regression: $x_i = i$; $x_i = 2^i$

✓ X

$y_i = \beta_0 + \beta_1 x_i + \sigma \varepsilon_i$ $\quad \hat{Y} = HY$

Estimating equations

- $g(\theta; X)$ is an unbiased estimating equation for θ if $\theta \in \mathbb{R}^p$

$$\mathrm{E}_\theta\{g(\theta; X)\} = 0, \quad \mathrm{E}_\theta\{g(\theta; X)g^T(\theta; X)\} < \infty$$

- given $X_1, \ldots, X_n$ i.i.d. with density $f(x; \theta)$, define the estimator $\tilde{\theta}_g$ by

$$\sum_{i=1}^{n} g(\tilde{\theta}_g; X_i) = 0$$

- then "under cond$^{\underline{n}}$s"

$$\sqrt{n}(\tilde{\theta}_g - \theta) \xrightarrow{d} N\{0, V(\theta)\}$$

- sandwich estimator $\quad V(\theta) = J^{-1}(\theta)I(\theta)J^{-1}(\theta)$

- $$J(\theta) = \mathrm{E}_\theta\{-g'(\theta; X)\}, \quad I(\theta) = \mathrm{E}_\theta\{g(\theta; X)g^T(\theta; X)\}$$

$g' = \frac{\partial}{\partial\theta} g(\theta; X)$

$\frac{\partial \ell^*(\theta; X)}{\partial \theta}$ is an example

$\hat{\theta}_{ML} - \theta \xrightarrow{d} N(0, I_1^{-1}(\theta))$

ML:

$J = I$

$\int \ell' f = 0$

If we using regression example

$x_1 = p \times 1$

$x_1^T = 1 \times p$

$\beta = p \times 1$

loss f =

$$\min_\beta \sum_{i=1}^n \rho(y_i - x_i^T \beta)$$

$g(\cdot) = \rho'(\cdot)$

i.e. $\frac{\partial}{\partial \beta} \sum \rho(y_i - x_i^T\beta) \Big|_{\hat\beta_{robust}} = 0$

$$= \sum_{i=1}^n \rho'(y_i - x_i^T\beta) x_i = 0$$

p eq's in p unk.

$y_i = x_i^T \beta + \varepsilon_i$

$A_n(\hat\beta_\rho - \beta) \xrightarrow{d} N(0, \sigma^2 J^{-1} I J^{-1})$

Robust regression

- limit theory derived from theory of estimating equations
- more generally, from the theory of model misspecification — MS 5.5
- true model $\quad X_1, \ldots, X_n$ i.i.d. $h(\cdot)$, say θ
- assumed model $\quad X_1, \ldots, X_n$ i.i.d. $f(\cdot; \theta)$

- maximum likelihood estimator — SM 4.6

$$\ell'(\hat\theta; \boldsymbol{X}) = 0$$

1. robust LS $\sum \rho'(y_i - x_i^T \tilde\theta) = 0$
2. est'g eq: $\sum g(y_i; \tilde\theta_g) = 0$
3. mis-spec'd mle

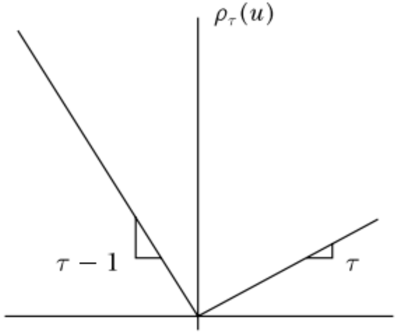
$$\hat\theta \xrightarrow{p} \arg\min_\theta \mathrm{E}_h \log\left\{\frac{h(x)}{f(x;\theta)}\right\}$$

relative entropy; K-L divergence

$$\sqrt{n}(\hat\theta - \theta) \xrightarrow{d} N\{0, J_h^{-1}(\theta_h) I_h(\theta_h) J_h^{-1}(\theta_h)\}$$

- forms the basis for GEE approach to longitudinal data

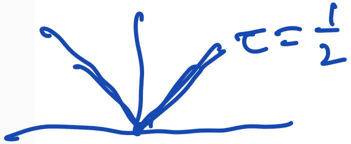
Figure 2
Quantile Regression ρ Function



$E(Y|X) = X\beta$ Lin.

$g\{E(Y|X)\} = X\beta$ GLM

$y_i = m(x_i) + \varepsilon_i$ $m(\cdot)$ "smooth"



$\rho_{\frac{1}{2}}(y_i - x_i^T\beta)$

$= |y_i - x_i^T\beta|$

$$\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta)$$

τ : τ^{th} quantile
$\in (0,1)$

Solution by linear programming; solution has approximately τ/n positive residuals

models med$(y_i | x_i)$

R package `quantreg`

Quantile regression

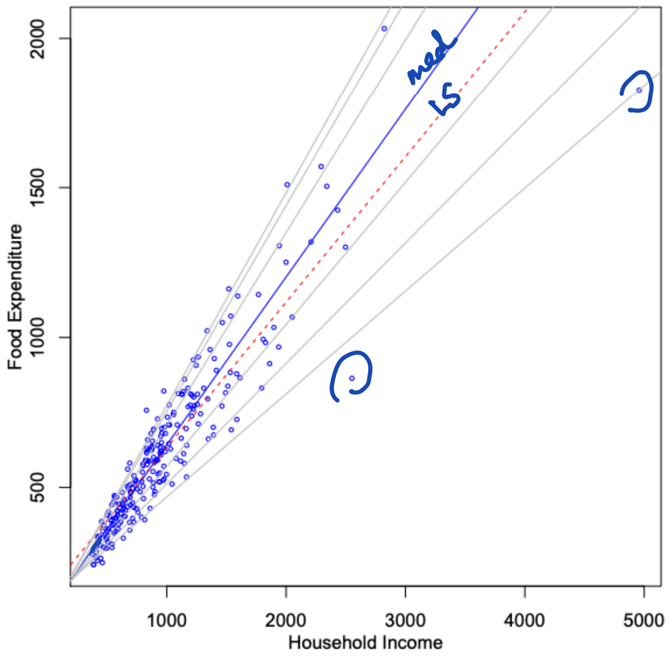


FIGURE 1. Scatterplot and Quantile Regression Fit of the Engel Food Expenditure Data: The plot shows a scatterplot of the Engel data on food expenditure vs household income for a sample of 235 19th century working class Belgian households. Superimposed on the plot are the $\{.05, .1, .25, .75, .90, .95\}$ quantile regression lines in gray, the median fit in solid black, and the least squares estimate of the conditional mean function as the dashed (red) line.

Likelihood asymptotics

$\hat{\theta} \pm 2\,\hat{se}$

```
> cover_prob
                1       5      10      50
Exact CI 0.95005 0.95148 0.95044 0.95101
q        0.94834 0.95741 0.95567 0.95159
s        0.71684 0.86771 0.90188 0.93992
r        0.93352 0.94634 0.94774 0.94939
r*       0.95165 0.95064 0.94940 0.94984
```

Figure 1: Simulation with 100000 times. Sample size $n = 1, 5, 10, 50$.

$$q = \frac{\hat{\theta} - \theta}{\hat{se}}$$

$$\alpha = P_n\{q \le c_\alpha\} \;\dot{=}\; P_n(Z \le c_\alpha)$$

$$\Rightarrow c_\alpha = z_\alpha$$

$$CB \left\{\theta : \frac{\hat{\theta} - \theta}{\hat{se}} \le z_\alpha \right\}$$

$X_1, \ldots, X_n$ i.i.d. $f(x;\theta) = \theta \exp(-\theta x)$

$$q = (\hat{\theta} - \theta) j^{1/2}(\hat{\theta}), \quad s = \ell'(\theta) j^{1/2}(\hat{\theta}), \quad r = \pm\sqrt{2\{\ell(\hat{\theta}) - \ell(\theta)\}}, \quad r^* = r + \frac{1}{r}\log\left(\frac{q}{r}\right)$$

Wald $\dot{\sim} N(0,1)$ Score $\dot{\sim} N(0,1)$ $\propto$ LRT $\dot{\sim} N(0,1)[1 + O(n^{-1/2})]$ $\dot{\sim} N(0,1)[1 + O(n^{-3/2})]$

First-order theory

ML theory

- central limit theorem: $Y_1, \ldots, Y_n$ i.i.d. $\mathrm{E}(Y_i) = \mu$, $\mathrm{var}(Y_i) = \sigma^2$, $< \infty$

$$\sqrt{n}(\bar{Y} - \mu)/\sigma \xrightarrow{d} N(0, 1)$$

- normal approximation

$$f(\bar{y}; \mu, \sigma^2) \doteq \frac{\sqrt{n}}{\sqrt{2\pi}\,\sigma} e^{-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2}$$

- saddlepoint approx

$$f_{\bar{Y}}(\bar{y}) \doteq \sqrt{\frac{n}{2\pi}}\{K_Y''(\hat{t})\}^{-\frac{1}{2}} e^{n\{K_Y(\hat{t}) - \hat{t}\bar{y}\}}$$

$$K_{Y_i}(t) = \log \underbrace{E(e^{tY_i})}_{\text{mgf}} = \log \int e^{ty} f_0(y)\,dy$$

cumulant gen fn

↑ true dist. of Y_i

$K_Y'(\hat{t}) = \bar{y} \qquad \hat{t} = \hat{t}(\bar{y}) \qquad K_Y''(\hat{t}) = K''(t)\big|_{t=\hat{t}}$

normal approx: $E\{e^{t(\bar{Y}-\mu)\sqrt{n}/\sigma}\} = M_n(t) \xrightarrow[n\to\infty]{} e^{t^2/2}$

$E\{e^{it\left(\frac{\sqrt{n}(\bar{Y}-\mu)}{\sigma}\right)}\} = \phi_n(t) \longrightarrow$ ~~d~~ chf for a N

$f_n(x) = \int e^{-itx}\phi_n(t)\,dt$

Fourier inversion

$$f_{\bar{Y}}(\bar{y}) = \sqrt{\frac{n}{2\pi}}\{K_Y''(\hat{t})\}^{-\frac{1}{2}} e^{n(K_Y(\hat{t}) - \hat{t}\bar{y})}$$

Ex. $Y_i \sim f(y_i;\theta) = e^{\theta y_i - c(\theta)} f_0(y_i)$ expl family form

$Y_1,\dots,Y_n$ iid

$$f(y;\theta) = e^{\theta\Sigma y_i - nc(\theta)}\prod_{i=1}^n f_0(y_i) \qquad [\theta\in\mathbb{R}]$$

$$K_{Y_i}(t) = \log E_\theta(e^{tY_i})$$

$$= \log\int e^{ty_i}e^{\theta y_i - c(\theta)} f_0(y_i)\,dy_i$$

$$= \log\left\{\int \underbrace{e^{(t+\theta)y_i} f_0(y_i)\,dy_i}\cdot e^{-c(\theta)}\right\}$$

$$= \log\left\{\underbrace{\int e^{(t+\theta)y_i - c(t+\theta)} f_0(y_i)\,dy_i}_{=1}\, e^{c(t+\theta)}\cdot e^{-c(\theta)}\right\}$$

$$= \log\{e^{c(t+\theta)-c(\theta)}\} = c(t+\theta) - c(\theta)$$

$$K_Y(t) = c(t+\theta) - c(\theta)$$

$$K_Y'(t) = c'(t+\theta) \qquad K_Y''(t) = c''(t+\theta)$$

$$K_Y'(\hat{t}) = \bar{y} = \underline{c'(\hat{t}+\theta)} \qquad L(\theta; y) = e^{\theta\Sigma y_i - nc(\theta)}$$

$$\ell(\theta; y) = \theta n\bar{y} - nc(\theta)$$

$$\Rightarrow \hat{t}+\theta = \hat{\theta} \qquad \underline{\hat{t} = \hat{\theta} - \theta} \qquad \ell'(\hat{\theta}) = 0 \Rightarrow \cancel{n}\bar{y} = \underline{\cancel{n}c'(\hat{\theta})}$$

$$K_Y''(\hat{t}) = c''(\hat{t}+\theta) = c''(\hat{\theta})$$

$$f_{\bar{Y}}(\bar{y};\theta) \doteq \sqrt{\frac{n}{2\pi}}\{c''(\hat{\theta})\}^{-1/2} \cdot e^{n c(\hat{\theta}) - nc(\theta) - n(\hat{\theta}-\theta)\bar{y}}$$

$$= \sqrt{\frac{n}{2\pi}}\{c''(\hat{\theta})\}^{-1/2} e^{n\{c(\hat{\theta}) - c(\theta) - (\hat{\theta}-\theta)c'(\hat{\theta})\}}$$

$c'(\hat{\theta}) = \bar{y}$

$\hat{\theta}$ is a fn of $\bar{y}$

$$\frac{d\bar{y}}{d\hat{\theta}} = c''(\hat{\theta})$$

change of variable

$$f_{\hat{\theta}}(\hat{\theta};\theta)\,d\hat{\theta} = f_{\bar{Y}}(\bar{y};\theta) \left|\frac{d\bar{y}}{d\hat{\theta}}\right| d\hat{\theta}$$

$$= \sqrt{\frac{n}{2\pi}}\{c''(\hat{\theta})\}^{\frac{1}{2}} \cdot e^{n\{\qquad\}}$$

$$\ell(\theta; y) = \theta n\bar{y} - nc(\theta) = \theta n c'(\hat{\theta}) - nc(\theta) = \ell(\theta; \hat{\theta})$$

$$= \sqrt{\frac{n}{2\pi}}\{c''(\hat{\theta})\}^{\frac{1}{2}} \cdot e^{\ell(\theta;\hat{\theta}) - \ell(\hat{\theta};\hat{\theta})}$$

$$f(\hat\theta;\theta) \doteq \sqrt{\frac{1}{2\pi}}\{j_n(\hat\theta)\}^{1/2} e^{\ell(\theta;\hat\theta)-\ell(\hat\theta;\hat\theta)}$$

$$j_n(\theta) = -\ell_n''(\theta) = -\frac{\partial^2 \ell}{\partial\theta^2}(\theta;y)$$

normal u § p.

$$\frac{1}{\sqrt{2\pi}}\{j_n(\hat\theta)\}^{\frac{1}{2}} e^{-\frac{j_n(\hat\theta)\{\theta-\hat\theta\}^2}{2}}$$

$$f(\hat\theta;\theta) \doteq \sqrt{\frac{1}{2\pi}}\{j_n(\hat\theta)\}^{1/2} e^{\ell(\theta;\hat\theta)-\ell(\hat\theta;\hat\theta)}$$

$$H_0: \theta=\theta_0 \qquad \Pr\{\hat\theta \geq \hat\theta^0;\ \theta_0\} = \int_{\hat\theta_0}^{\infty} f(\hat\theta;\theta_0)\,d\hat\theta$$

<u>obs'd p-value</u>

$$1-p_{val}^{obs} = \int_{-\infty}^{\hat\theta^0} f(\hat\theta;\theta)\,d\hat\theta \doteq \int_{-\infty}^{\hat\theta^0}\sqrt{\frac{1}{2\pi}}\,(\hat j)^{\frac{1}{2}} e^{\ell(\theta;\hat\theta)-\ell(\hat\theta;\hat\theta)}\,d\hat\theta$$

$$\ell(\theta;\hat\theta)-\ell(\hat\theta;\hat\theta) = -\frac{1}{2}r^2 \qquad r = r(\hat\theta;\theta)$$

c. of. var $\hat\theta \to r$

$$-r\,dr = \{\ell_{;\hat\theta}(\theta;\hat\theta) - \ell_{;\hat\theta}(\hat\theta;\hat\theta) - \ell_{;\hat\theta}(\hat\theta;\hat\theta)\}\,d\hat\theta$$

$$\ell(\theta;\hat\theta) = n\,\theta c'(\hat\theta) - nc(\theta)$$

$$\frac{\partial \ell(\theta;\hat\theta)}{\partial\hat\theta} = n\theta c''(\hat\theta) \qquad \cancel{\frac{\partial\ell(\hat\theta;\hat\theta)}{\partial\hat\theta}} = \ell_{;\hat\theta}(\hat\theta;\hat\theta) = n\hat\theta c''(\hat\theta)$$

$$= \int_{-\infty}^{r^0}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2}\cdot \hat j^{1/2}\cdot\frac{r}{n\hat\theta c''(\hat\theta)-n\theta c''(\hat\theta)}\,dr \qquad nc''(\hat\theta) = j_n(\hat\theta)$$

$$= \int_{-\infty}^{r^0}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2}\frac{r}{j^{1/2}(\hat\theta)(\hat\theta-\theta)}\,dr$$

$$= \int_{-\infty}^{r_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} \frac{r}{q}\, dr \qquad r^* = r + \frac{1}{r}\log\frac{q}{r}$$

$$= \int_{-\infty}^{r_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2 + \log\frac{r}{q}}\, dr \qquad r^* \sim N(0,1)$$

$$= \int_{-\infty}^{r_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\underbrace{\{r^2 \cancel{+} 2\log\left(\frac{r}{q}\right) + \frac{1}{r^2}\log^2\left(\frac{r}{q}\right)\}}_{\left(r + \frac{1}{r}\log\frac{q}{r}\right)^2}} e^{\frac{1}{2}r^{-2}\log^2\left(\frac{r}{q}\right)}\, dr$$

$$= \int_{\infty}^{r_0} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}r^{*2}}} \cdot e^{\frac{1}{2}r^{-2}\log^2\left(\frac{q}{r}\right)}\, dr$$

$$e^{\frac{1}{2}r^{-2}\log^2\left(\frac{q}{r}\right)} \doteq \left(1 + \frac{1}{2}r^{-2}\log^2\frac{q}{r}\right)$$

$$\underline{\varphi(r^*)} + \boxed{\int \varphi(r^*) \cdot \frac{1}{2}r^{-2}\log^2\frac{q}{r}\, dr}$$

? remainder ≈ 0 ?

$$r^* = r + \frac{1}{r}\log\left(\frac{q}{r}\right)$$

$$r = \pm\sqrt{2\{\ell(\hat\theta) - \ell(\theta)\}} \qquad \text{LRT}$$

$$q = (\hat\theta - \theta)\, j^{1/2}(\hat\theta) \qquad \text{Wald}$$

$$r = q + \frac{A}{\sqrt{n}}q^2 + \frac{B}{n}q^3 + O(n^{-3/2})$$

$\ell(\theta) = \ell(\hat\theta)$

$+ \cancel{(\theta-\hat\theta)\ell'(\hat\theta)}$

$+ \frac{1}{2}(\theta-\hat\theta)^2\ell''(\hat\theta)$

$+ \dots \quad q^2$

$$q = r + \frac{a}{\sqrt{n}} r^2 + \frac{b}{n} r^3 + O(n^{-3/2})$$

$$\frac{1}{r} \log\left(\frac{q}{r}\right) \doteq \log\left(1 + \frac{a}{\sqrt{n}} r + \frac{b}{n} r^2\right) \qquad \log(1+x) = x - \frac{1}{2}x^2 + \ldots$$

$$= \frac{1}{r}\left(\frac{a}{\sqrt{n}} r + \frac{b}{n} r^2 - \frac{1}{2}\frac{a^2}{2n} r^2\right) +$$

$$= \frac{a}{\sqrt{n}} + \left(\frac{b}{n} - \frac{1}{2}\frac{a^2}{n}\right) r$$

$$(\;\downarrow\;)^2 = \frac{a}{n} + O(n^{-3/2})$$

$$\int_{-\infty}^{r_0} \varphi(r^*)\left(1 + \frac{a}{n}\right) dr \Rightarrow \text{a tiny bit } \underline{\text{more}}$$

$$f(\hat\theta; \theta) \doteq \text{saddlepoint approx}^{n} \text{ to } \bar{Y}$$

$$f(r^*; \theta) \doteq N(0,1) \qquad r^* = r + \frac{1}{r} \log \frac{q}{r}$$

$$f(\hat\theta; \theta) \doteq \frac{1}{\sqrt{2\pi}} j^{1/2}(\hat\theta)\, e^{\ell(\theta) - \ell(\hat\theta)} \qquad \int f(\hat\theta;\theta)\, d\hat\theta = 1 + \frac{d}{n}$$

$$f(\hat\theta; \theta) \doteq c \cdot j^{1/2}(\hat\theta)\, e^{\ell(\theta) - \ell(\hat\theta)} \qquad (p^*\text{-formula})$$

$$\cancel{\Phi(r^*)} \quad r^* \sim N(0,1) \qquad (r^*\text{-approx}^n)$$

$$\pi(\theta|y) = \frac{e^{\ell(\theta)}\pi(\theta)}{\int e^{\ell(\theta)}\pi(\theta)\,d\theta}$$

$\ell(\theta) = \ell(\hat\theta) + (\theta-\hat\theta)\ell'(\hat\theta)\dots + \frac{1}{2}(\theta-\hat\theta)^2\ell''(\hat\theta)$

Laplace app$^{\underline{\times}}$

$$\doteq e^{\ell(\theta)-\ell(\hat\theta)} \frac{\{j(\hat\theta)\}^{1/2}}{\sqrt{2\pi}} \frac{\pi(\theta)}{\pi(\hat\theta)}$$

$$\int^{\theta^*} \pi(\theta|y)\,d\theta = \int \qquad d\theta$$

$$r^*_{Bayes} \; \dot{\sim} \; N(0,1) = r + \frac{1}{r}\log\frac{s_{Bayes}}{r}$$

$$s = \ell'(\theta)\{j(\hat\theta)\}^{-1/2} \cdot \frac{\pi(\theta)}{\pi(\hat\theta)}$$

$n\bar{X}_n = n\bar{D}$ (iii) $\bar{X} = \bar{D}$

$$T_n = \frac{1}{n\bar{x}_n^2}\sum D_i^2$$

$$= \frac{1}{n\bar{D}^2}\sum D_i^2$$

$$= \frac{1}{n}\sum_{i=1}^n \left(\frac{D_i}{\bar{D}}\right)^2$$

$$\sqrt{n}(T_n - 2) \xrightarrow{d} N(0, 20) \quad (4?)$$

$$Y_i = \left(\frac{D_i}{\bar{D}}\right)^2 \qquad E(Y_i) = ?$$

$var(Y_i) = ?$

$D_1 \dots D_n \overset{1-1}{\longleftrightarrow} \left(\frac{D_1}{\bar D}, \dots \frac{D_n}{\bar D}, \bar D\right)$

$cov(D_i, \bar D) = \frac{1}{n}\sigma_D^2 \quad = cov\left(D_i, \frac{1}{n}(D_1 + \dots + D_n)\right)$

$= cov(D_i, D_i)$

(wlog $\lambda = 1$)

$\underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} D_i^2}_{\downarrow N(\mu_i^2 \dots)} / \sqrt{n}\bar D_n^2 \xrightarrow{p} ?$

$\cancel{Z}_n = \frac{1}{n}\sum D_i^2$

$W_n = \frac{1}{\bar D^2}$

$\cancel{Z}_i$ iid (μ_Z, σ_Z^2) $\quad \cancel{\sqrt{n}}\left(\frac{\sum Z_i}{\sqrt{n}} - \mu\right) \to N(0, \sigma_Z^2)$