Lehmann & Romano, TSH Ch. 3

- Setup: define a test function $\phi(y)$ from $\mathcal{Y}$ to $[0, 1]$
- $\phi(Y) = \Pr(Y \in \mathcal{R})$
- if $\phi(y) = 1$ then $y \in \mathcal{R}$, if 0, $y \notin \mathcal{R}$
- allows for the possibility of randomized tests

- if $Y \sim f(y; \theta)$, then
- $E_\theta \phi(Y) = \int \phi(y) f(y; \theta) dy$ = probability of rejection
- under $H_0 : \theta \in \Theta_0$, this is the size of the test, or type I error
- under $H_1 : \theta \in \Theta_1$, this is the power of the test

- Goal: maximize

$$\beta_\phi(\theta) = E_\theta \phi(Y) \quad \forall \theta \in \Theta_1,$$

subject to

$$E_\theta \phi(Y) \leq \alpha, \quad \forall \theta \in \Theta_0$$

Neyman-Pearson lemma

- Suppose Θ_0 is the point θ_0, and similarly for Θ_1
- Assume the existence of densities f_0 and f_1 with respect to the same measure μ

1. Given $0 \leq \alpha \leq 1$, there exists a test function ϕ and a constant k such that
$$E_0\phi(Y) = \alpha \tag{1}$$
and
$$\phi(y) = \begin{cases} 1 & \text{when} \quad f_1(y) > kf_0(y), \\ 0 & \text{when} \quad f_1(y) < kf_0(y). \end{cases} \tag{2}$$
2. If a test satisfies (1) and (2) for some k, then it is most powerful for testing f_0 against f_1 at level α
3. If ϕ is most powerful at level α for testing f_0 against f_1, then for some k it satisfies (2), a.e. μ, and satisfies (1) unless there exists a test of size $< \alpha$ and with power 1.

Proof 1.

- trivial for $\alpha = 0$ and $\alpha = 1$ allow $k = \infty$
- 1. define
  $\alpha(c) = \Pr_0\{f_1(Y) > cf_0(Y)\} = \Pr\{f_1(Y)/f)0(Y) > c\}$.
- $1 - \alpha(c)$ is a cumulative distribution function
- so $\alpha(c)$ is non-increasing, right-continuous,
  $\alpha(-\infty) = 1, \alpha(\infty) = 0$

- Given $0 < \alpha < 1$, let c_0 be such that $\alpha(c_0) \le \alpha \le \alpha(c_0^-)$
- $$\phi(y) = \begin{cases} 1 & \text{when} \quad f_1(y) > c_0 f_0(y) \\ \frac{\alpha - \alpha(c_0)}{\alpha(c_0^-) - \alpha(c_0)} & \text{when} \quad f_1(y) = c_0 f_0(y) \\ 0 & \text{when} \quad f_1(y) < c_0 f_0(y) \end{cases}$$
- $$E_0\phi(Y) = \Pr_0\left\{\frac{f_1(Y)}{f_0(Y)}\right\} +$$

... proof 2.

- Suppose ϕ is a test satisfying (1) and (2), and that ϕ^* is another test with $E_0\phi^*(Y) \le \alpha$.
- Denote by S^+ and S^- the sets in $\mathcal{Y}$ where $\phi(y) - \phi^*(y) > 0$ and < 0.
- In S^+, $\phi(y) > 0$ so $f_1(y) \ge kf_0(y)$, and
- $$\int (\phi - \phi^*)(f_1 - kf_0)d\mu = \int_{S^+ \cup S^-} (\phi - \phi^*)(f_1 - kf_0)d\mu \ge 0$$
- difference in power:

$$\int (\phi - \phi^*) f_1 d\mu \ge k \int (\phi - \phi^*) f_0 d\mu \ge 0$$

... proof 3.

- Let ϕ^* be MP level α, and ϕ satisfy (1) and (2)
- On $S^+ \cup S^-$, ϕ and ϕ^* differ. Let $S = S^+ \cup S^- \cap \{y : f_1(y) \neq kf_0(y)\}$, and assume $\mu(S) > 0$
- $$\int_{S^+ \cup S^-} (\phi - \phi^*)(f_1 - kf_0) d\mu = \int_S (\phi - \phi^*)(f_1 - kf_0) d\mu > 0$$
- implies ϕ is more powerful than ϕ^*
- contradiction, hence $\mu(S) = 0$

- if ϕ^* had size $< \alpha$ and power < 1, could add points to rejection region until either $E_0\phi^*(Y) = \alpha$ or $E_1\phi^*(Y) = 1$

- test is unique if $\{y : f_1(y) = kf_0(y)\}$ has measure 0

Comments

- discreteness: e.g. $Y \sim \text{Bin}(n, p)$
- MP test has rejection region $\mathcal{R}$ determined by $\{y > d_\alpha\}$
- not all values of α attainable: e.g. CH Example 4.9
  $Y \sim \text{Poisson}(\mu)$
- $H_0 : \mu = 1, \quad H_1 : \mu = \mu_1 > 1$, MP test $Y \geq d_\alpha$

Table : attained significance levels

y	$\Pr(Y > y; \mu = 1)$	y	$\Pr(Y > y; \mu = 1)$
0	1	4	0.0189
1	0.632	5	0.0037
2	0.264	6	0.0006
3	0.080	$\vdots$	$\vdots$

- if critical regions are *nested*, i.e. $\mathcal{R}_{\alpha_1} \subset \mathcal{R}_{\alpha_2}, \alpha_1 < \alpha_2$, then
  $p_{obs} = \inf(\alpha; y_{obs} \in \mathcal{R}_\alpha)$
- asymmetry:
  $Y \sim N(\mu, 1), H_0 : \mu = 0, H_1 : \mu = 10, \quad y_{obs} = 3$

Bayesian testing

see CH Example 10.12

- simple H_0, simple H_1:

$$\frac{\Pr(H_0 \mid y)}{\Pr(H_1 \mid y)} = \frac{\Pr(H_0)}{\Pr(H_1)} \frac{f_0(y)}{f_1(y)}$$

- similarly, with $H_1, \dots H_k$ potential alternatives

$$\frac{\Pr(H_0 \mid y)}{\Pr(H_0^c \mid y)} = \frac{\Pr(H_0) f_0(y)}{\Sigma_j \Pr(H_j) f_j(y)}$$

- sharp null hypothesis: $H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0$

$$\frac{\Pr(H_0 \mid y)}{\Pr(H_0^c \mid y)} = \frac{\pi_0}{(1-\pi_0)} \frac{f(y;\theta_0)}{\int \pi_1(\theta) f(y;\theta) d\theta}$$

- nuisance parameters

$$\frac{\Pr(H_0 \mid y)}{\Pr(H_0^c \mid y)} = \frac{\pi_0}{(1-\pi_0)} \frac{\pi(\lambda \mid h_0) f(y \mid \psi_0, \lambda) d\lambda}{\int \int \pi(\psi, \lambda \mid H_1) f(y \mid \psi, \lambda) d\psi d\lambda}$$

... testing

- Bayes factor $B_{10} = \dfrac{\Pr(y \mid H_1)}{\Pr(y \mid H_0)}$
- typically $\Pr(y \mid h_i) = \int f(y \mid H_i, \theta_i)\pi(\theta_i \mid H_i) d\theta_i, \quad i = 0, 1$

11.2 · Inference

Table 11.3
Interpretation of Bayes factor B_{10} in favour of H_1 over H_0. Since $B_{10} = B_{01}^{-1}$, negating the values of $2 \log B_{10}$ gives the evidence against H_1.

B_{10}	$2 \log B_{10}$	Evidence against H_0
1–3	0–2	Hardly worth a mention
3–20	2–6	Positive
20–150	6–10	Strong
> 150	> 10	Very strong

SM Ch. 11.2

- cannot be computed with improper priors

Nature, PNAS, AoS articles by Johnson

- developed an ‘objective’ Bayesian test for comparison to *p*-values
- “A *p*-value of 0.05 or less corresponds to Bayes factors of between 3 and 5, which are consider weak evidence to support a finding”
- “He advocates for scientists to use more stringent *p*-values of 0.005 or less”
- see also CH Example 10.12 and SM Example 11.15
- emphasis on point hypotheses drives most of these anomalous results
- e.g. $\Pr(\theta > 0 \mid y)$