U.S. FOOD & DRUG ADMINISTRATION

Weighted Posterior Odds: A Data Summary for Decision Making

GENE PENNELLO, PhD

Division of Imaging Diagnostics & Software Reliability
U.S. Food and Drug Administration
Center for Devices and Radiological Health
Office of Science and Engineering Laboratories

Bayesian Biostatistics 2024, Rockville, MD – USA
Thursday, 10/24/2024, 11:55 AM - 12:15 PM

Disclosure / Disclaimer

- The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

- This is a contribution of the U.S. Food and Drug Administration is not subject to copyright.

Outline

- Bayesian model, loss function for testing a 1-sided hypothesis
  - Bayes rule critical region: weighted posterior odds is the focal point
  - Normal data, linear loss as a special case
  - Numerical examples
- Bayes model, loss function for testing a 1-sided comparison of means
  - Bayes rule critical region: weighted posterior odds is the focal point
  - Normal data, linear loss as a special case
  - Numerical examples
- Linear loss in testing $\Longleftrightarrow$ squared-error loss in estimation connection
- Discussion

Bayesian Model

Data

$$\underline{\boldsymbol{y}} = (y_1, y_2, \dots y_n)$$

$$y_i \overset{iid}{\sim} f(y|\mu, \gamma)$$

Hypothesis Test

$$H: \mu \leq \mu_0$$

$$A: \mu > \mu_0$$

Loss Model

Loss Function $L(a, \mu) = L_a(\mu)$

Decision $a = I\left(\underline{y} \in R\right)$

Critical Region $R = \left\{\underline{\boldsymbol{y}}: \text{accept } A: \mu > \mu_0 \text{ as true}\right\}$

Type 1 Loss $L_1(\mu) = \boldsymbol{k} l(\mu_0, \mu) I(\mu \leq \mu_0)$

Type 2 Loss $L_0(\mu) = \; l(\mu, \mu_0) I(\mu > \mu_0)$

Distance Function $l(u, v) \begin{cases} = 0, u \leq v \\ > 0, u > v \end{cases}$

$\boldsymbol{k}$ = type 1 to type 2 error loss (or seriousness) ratio

Bayes Rule Critical Region

$$\tilde{R} = \left\{ \underline{\boldsymbol{y}} : W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k \right\}$$

where

$$W\left(\underline{\boldsymbol{y}}, \mu_0\right) = \frac{\int_{\mu_0}^{\infty} l(\mu,\mu_0)\pi\left(\mu|\underline{\boldsymbol{y}}\right)d\mu}{\int_{-\infty}^{\mu_0} l(\mu_0,\mu)\pi\left(\mu|\underline{\boldsymbol{y}}\right)d\mu} =$$ **weighted posterior odds of $\boldsymbol{A}$**

$\pi\left(\mu|\underline{\boldsymbol{y}}\right) =$ posterior of $\mu|\underline{y}$

Bayes Rule Critical Region

$$\tilde{R} = \left\{\underline{\boldsymbol{y}}: W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k\right\}$$

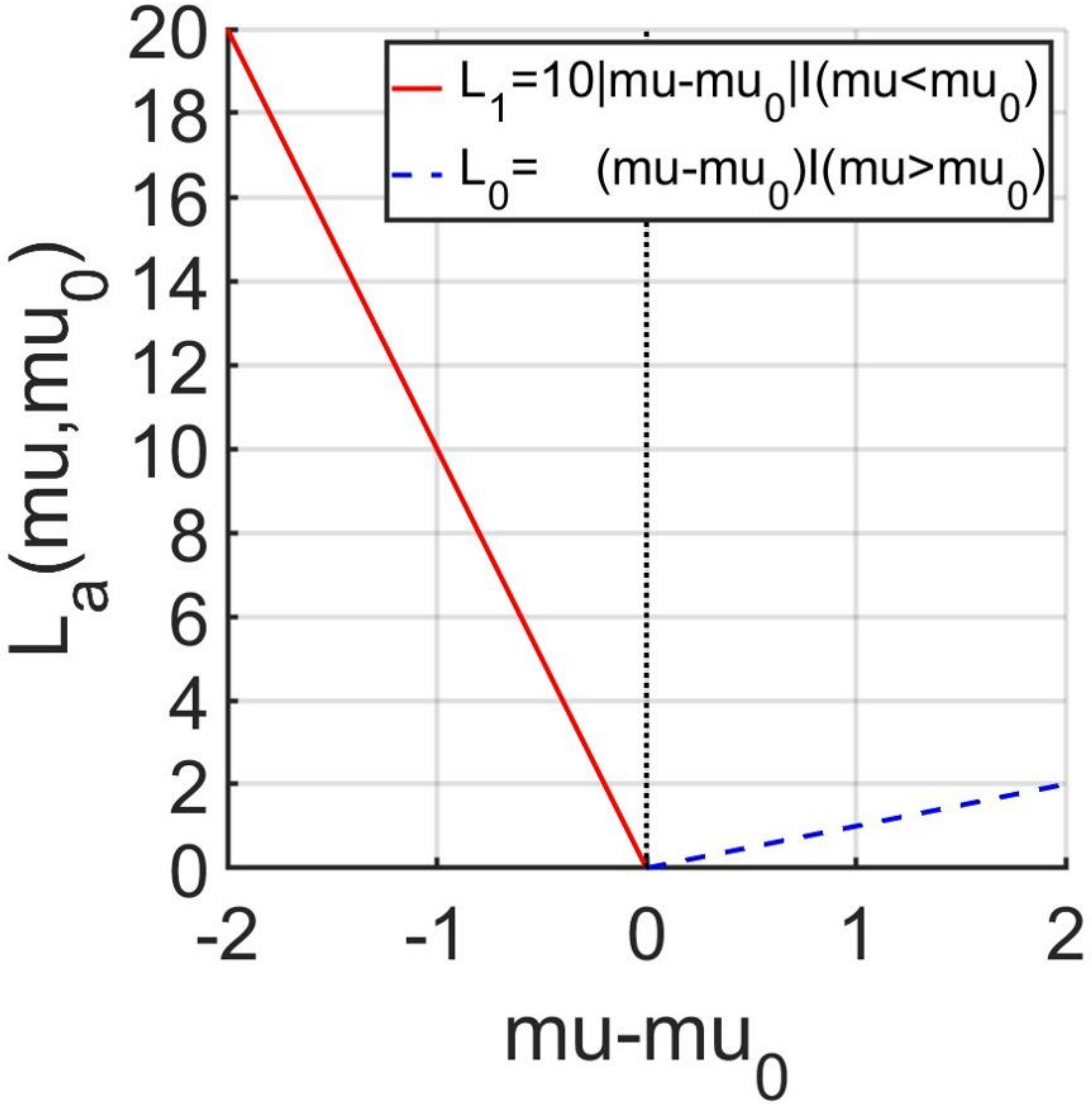
Note $W\left(\underline{\boldsymbol{y}}, \mu_0\right) = \sup_k \left\{k: W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k\right\}$

= supremum of the values of k for which Bayes rule accepts $A: \mu > \mu_0$.

Interpretation

- If $W\left(\underline{\mathbf{y}}, \mu_0\right) = \mathbf{100}$, then the data support accepting A when
  - **type 1 error is no more than 100 times more serious than type 2 error**, where
  - type 1 error = falsely accepting $A: \mu > \mu_0$
  - type 2 error = falsely not accepting $A: \mu > \mu_0$
- $W\left(\underline{\mathbf{y}}, \mu_0\right)$ measures the evidence against H and in favor of A on the scale of
  - **the trade-off between the two potential decision errors of**
  - falsely accepting A (type 1) and falsely not accepting A (type 2).
- In contrast, the $\boldsymbol{p}$ **value** measures evidence against H, but not on the scale of the trade-off between the potential decision errors.
- Thus, for decision making, the p value is arguably less interpretable than $W\left(\underline{\mathbf{y}}, \mu_0\right)$.

Linear Loss Function: $l(u, v) = u - v$



Bayes Rule Critical Region under Linear Loss

$$\tilde{R} = \left\{ \underline{\boldsymbol{y}} : W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k \right\}$$

where

$$W\left(\underline{\boldsymbol{y}}, \mu_0\right) = \frac{\int_{\mu_0}^{\infty} (\mu - \mu_0)\pi\left(\mu | \underline{\boldsymbol{y}}\right) d\mu}{\int_{-\infty}^{\mu_0} (\mu_0 - \mu)\pi\left(\mu | \underline{\boldsymbol{y}}\right) d\mu} = \textbf{weighted posterior odds of } \boldsymbol{A}$$

$$\pi\left(\mu | \underline{\boldsymbol{y}}\right) = \text{posterior of } \mu | \underline{\boldsymbol{y}}$$

Normal Data; Diffuse Prior

Data

$$\underline{\boldsymbol{y}} = (y_1, y_2, \ldots y_n)\; y_i \overset{iid}{\sim} N(\mu, \tau^{-1}), \tau = \sigma^{-2}$$

$$\bar{y} = n^{-1}\Sigma y_i \quad \sim N\left(\mu, \tau_{\bar{y}}^{-1}\right), \tau_{\bar{y}} = n\tau$$

$$s^2 = f^{-1}\Sigma(y_i - \bar{y})^2 \sim \Gamma\left(\frac{f}{2}, \frac{f\tau}{2}\right), f = n - 1$$

Prior (Box-Tiao)

$$\pi(\mu, \tau) \propto \tau^{-1}$$

Normal Data; Box-Tiao Prior

Posterior

$$\pi\left(\mu,\tau|\underline{\boldsymbol{y}}\right)=\pi\left(\mu|\tau,\underline{\boldsymbol{y}}\right)\pi\left(\tau|\underline{\boldsymbol{y}}\right)$$

$$\mu|\tau,\underline{\boldsymbol{y}}\sim N\left(\overline{y},\tau_{\overline{y}}^{-1}\right)$$

$$\tau|\underline{\boldsymbol{y}}\sim\Gamma\left(\frac{f}{2},\frac{fs^{2}}{2}\right)$$

Bayes Rule Critical Region

$$\tilde{R} = \left\{ \underline{\boldsymbol{y}} : W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k \right\}$$

$$W\left(\underline{\boldsymbol{y}}, \mu_0\right) = M(t, f)/M(-t, f)$$

$$t = \sqrt{n}(\overline{y} - \mu_0)/s$$

$$f = n - 1$$

$$M(t, f) = (f - 1)(t^2 + f)g(z, f) + tG(z, f)$$

$g(\bullet, \nu)$ = density of Student-t with ν dof

$G(\bullet, \nu)$ = cumulative dist’n of Student-t with ν dof

Bayes Rule Critical Region, τ known

$$\tilde{R} = \left\{ \underline{\boldsymbol{y}} : W\left(\underline{\boldsymbol{y}}, \mu_0\right) > k \right\}$$

$$W\left(\underline{\boldsymbol{y}}, \mu_0\right) = M(z)/M(-z)$$

$$z = \sqrt{n}(\bar{y} - \mu_0)/\sigma$$

$$M(z) = \varphi(z) + z\Phi(z) = \lim_{f \to \infty} M(t, f)$$

$\varphi(\bullet)$ = density of $N(0,1)$ variable

$\Phi(\bullet)$ = cumulative dist'n of $N(0,1)$ variable

Interval Estimation

- $\boldsymbol{k}$**CI**$= (\tilde{\boldsymbol{\mu}}_L, \tilde{\boldsymbol{\mu}}_U)$ = values of μ_0 for which neither $A: \mu > \mu_0$ nor $A': \mu < \mu_0$ can be accepted based on the hypothesis tests of $H: \mu \leq \mu_0$ and $H': \mu \geq \mu_0$.

- When $k > 1$ is the same for both hypothesis tests,

$$\tilde{\mu}_L \text{ satisfies } W\left(\underline{\boldsymbol{y}}, \tilde{\mu}_L\right) = k$$

$$\tilde{\mu}_U \text{ satisfies } W\left(\underline{\boldsymbol{y}}, \tilde{\mu}_U\right) = k^{-1}$$

- For normal data, linear losses, and diffuse Box-Tiao prior,

$$k\text{CI} = \bar{y} \pm t_k s/\sqrt{n}, \text{ where } t_k \text{ satisfies } \frac{M(t_k, f)}{M(-t_k, f)} = k$$

Numerical Examples

$\bar{y}$	s^2	f	t	p val	$W(\underline{y}, \mu_0)$	95% CI	kCI
1.66	1	9	2.08	0.068	100.0	0.94,2.37	1.00,2.31
2.26	1	9	2.26	0.050	146.4	1.00,2.43	1.06,2.37
1.54	1	∞	1.72	0.085	100.0	0.92,2.16	1.00,2.09
1.62	1	∞	1.96	0.050	208.5	1.00,2.24	1.08,2.16

$\bar{y} = n^{-1}\Sigma y_i, s^2 = f^{-1}\Sigma(y_i - \bar{y})^2, f = n - 1$ or $\infty, t = \sqrt{n}(\bar{y} - \mu_0)/s,$

kCI$= \bar{y} \pm t_k s/\sqrt{n}$, where t_k satisfies $\frac{M(t_k, f)}{M(-t_k, f)} = k$

Two-Sample Comparison of Means

FDA

Data

$$y_{gi} \sim N\left(\mu_g, \tau^{-1}\right), \tau = \sigma^{-2}$$

Prior (Box-Tiao)

$$\mu_g \sim N\left(\mu_0, \tau_\mu^{-1}\right)$$

$$\pi\left(\mu_0, \mu_1 \tau, \tau_\mu\right) \propto \tau\left(\tau^{-1} + 2\tau_\mu^{-1}\right)^{-1}$$

$$= \tau_\mu\left(\tau_\mu + 2\tau\right)^{-1}$$

Hypothesis Test

$$\delta = \mu_2 - \mu_1$$

$$H: \delta \leq \delta_0$$

$$A: \delta > \delta_0$$

Bayes Rule Critical Region, Linear Loss

$$\tilde{R} = \{\underline{\boldsymbol{y}} : W(\underline{\boldsymbol{y}}; \delta_0) > k\}$$

$$W(\underline{\boldsymbol{y}}; \delta_0) = W(t, f; \delta_0)$$

$$t = (\bar{d} - \delta_0)/s_{\bar{d}}$$

$$s_{\bar{d}}^2 = 2s^2/n$$

$$s^2 = f^{-1}\Sigma\Sigma(y_{ai} - \bar{y}_a)^2$$

$$f = n - 1$$

Weighted Posterior Odds, Linear Loss

$$W(t, f; \delta_0) = \sup_k \{k: W(t, f; \delta_0) > k\}$$

$=$ supremum of the values of k at which $A: \delta > \delta_0$ is accepted by Bayes rule.

- If $W(t, f; \delta_0) = \mathbf{100}$, then the data support accepting A when
  - **type 1 error is no more than <u>100 times more serious</u> than type 2 error**, where
  - type 1 error = falsely accepting $A: \delta > \delta_0$
  - type 2 error = falsely not accepting $A: \delta > \delta_0$

Numerical Examples

$t_\gamma(f) =$ γth quantile of Student-t with f dof

$$t_{0.975}(\infty) = 1.96\text{: } W(1.96, \infty; \delta_0) \cong 100$$

$$t_{0.95}(\infty) = 1.645\ W(1.645, \infty; \delta_0) \cong 50$$

$$t_{0.995}(\infty) = 2.576\ W(2.576, \infty; \delta_0) \cong 500$$

Interval Estimation

- $\boldsymbol{k}\mathbf{CI} = \left(\tilde{\delta}_L\,, \tilde{\delta}_U\right) =$ values of μ_0 for which neither $A\!:\delta > \delta_0$ nor $A'\!:\delta < \delta_0$ can be accepted based on the hypothesis tests of $H\!:\delta \leq \delta_0$ and $H'\!:\delta \geq \delta_0$.

- When $k > 1$ is the same for both hypothesis tests,

$$\tilde{\delta}_L \text{ satisfies } W\left(t, f; \tilde{\delta}_L\right) = k$$

$$\tilde{\delta}_U \text{ satisfies } W\left(t, f; \tilde{\delta}_U\right) = k^{-1}$$

- For normal data, linear losses, and diffuse Box-Tiao prior,

$$k\mathrm{CI} = \bar{d} \pm t_k\, s\sqrt{2/n}, \text{ where } t_k \text{ satisfies } W(t_k, f; \delta_0) = k$$

Squared-Error Loss for Estimation

- **Result**. Squared-error loss for estimation and linear loss for testing yield the same interval estimate $(\tilde{\delta}_L, \tilde{\delta}_U)$:
  - $\tilde{\delta}_L$ satisfies $W(\underline{\boldsymbol{y}}; \tilde{\delta}_L) = k$
  - $\tilde{\delta}_U$ satisfies $W(\underline{\boldsymbol{y}}; \tilde{\delta}_U) = k^{-1}$

Squared Error Loss $L(e,\delta) = \begin{cases} k(e-\delta)^2, e \geq \delta \\ (\delta - e)^2, e < \delta \end{cases}$

Posterior Expected Loss $E[L(e,\delta)] = \int_{-\infty}^{e} k(e-\delta)^2 \pi(\delta|\underline{y})\, d\delta + \int_{e}^{\infty} (\delta - e)^2 \pi(\delta|\underline{y})\, d\delta$

$$\frac{\partial}{\partial e} E[L(e,\delta)] \overset{\text{Leibnitz}}{=} 2k \int_{-\infty}^{e} (e-\delta)\pi(\delta|\underline{y})\, d\delta - 2\int_{e}^{\infty} (\delta - e)\pi(\delta|\underline{y})\, d\delta \overset{\text{set}}{=} 0$$

$$\therefore\ k = \int_{e}^{\infty} (\delta - e)\pi(\delta|\underline{y})\, d\delta \Big/ \int_{-\infty}^{e} (e-\delta)\pi(\delta|\underline{y})\, d\delta \equiv W(\underline{\boldsymbol{y}}; \delta_0)$$

Dixon DO, Duncan DB. Minimum Bayes Risk *t* -Intervals for Multiple Comparisons, *JASA* 1975; 70 (352), 822-831.
Dixon DO. Interval Estimates Derived from Bayes Testing Rules. *JASA* 1976; 71 (354), 406-408.

Discussion

- **Summaries of evidence for a hypothesis include**
  - the $\boldsymbol{p}$ **value**, the smallest allowed type 1 error level for accepting a hypothesis,
  - **weighted posterior odds (WPO)**, the largest allowed type-1-to-type-2-error-seriousness ratio (k) for accepting a hypothesis.
- **Arguably, WPO is of more interest to decision makers than the $\boldsymbol{p}$ value.**
  - Decision makers make value judgments about the consequent benefits and harms of actions that would be taken if a hypothesis were accepted as true or not.
  - Likewise, WPO summarizes evidence for a hypothesis in terms of relative consequences of decision errors, where "consequence"= "seriousness", "importance", "harm", "risk", "loss"...
- **For stakeholders** (e.g., regulator, payer, provider, policy maker, manufacturer, doctor, patient), **the observed WPO in a study**
  - may be compared with their own value judgment of k.
  - should facilitate better discussion as to whether to accept a hypothesis as true or not.

Discussion *(continued)*

- **Proposal:** Set preliminary value of k by FDA regulatory pathway, e.g.,

Product	k_1	k_2	k_3	k_4
Drug / Biologic	Full approval	Accelerated approval	Orphan	
		Biosimilar		
Medical Device	Class III	Class II	Humanitarian	Class I
	Breakthrough	Lab developed test	EUA	Wellness
			Wearable	

$$k_1 > k_2 > k_3 > k_4 \gg 1$$

Discussion (*continued*)

- **The k-ratio method** was developed by **David B. Duncan** and his students for Bayes rule **multiple comparisons of means (MCM)** problems assuming
  - (1) **additive linear losses** of the component comparisons,
  - (2) **prior exchangeability** of the means

- **MCM Problems**
  - Comparisons of Means in a 1-Way Array (Waller)
    - $\mathrm{W}\left(t_{ij,}, F, q, f; \delta_0\right) =$ WPO for $\delta_{ij} > \delta_0$
  - Comparisons of Treatments with a Control (Brant)
    - $\mathrm{W}(t_{i0}, F_T, F_G, q_T, q_G, f; \delta_0) =$ WPO for $\delta_{i0} > \delta_0$ $(q_G = 1)$
  - Largest Mean Problem (Ranking and Selection) (Bland)
  - Comparisons of Means in a 2-Way Array (Pennello)
    - $\mathrm{W}\left(t_{ij,k}, t_{ij\bullet}, F_A, F_B, F_C, q_A, q_B, q_C, f; \delta_0\right) =$ WPO for $\delta_{ij,k} > \delta_0$

Discussion *(continued)*

Extra Results. If distance function $l(u, v)$ is constant when $u > v$ and 0 otherwise, then

(1) $W\left(\underline{y}, \mu_0\right) = \boldsymbol{\pi_A}(\mathbf{1} - \boldsymbol{\pi_A})^{-\mathbf{1}}$ is the posterior odds of $A\colon \mu > \mu_0$, where $\boldsymbol{\pi_A} = \mathbf{Pr}\left(\mu > \mu_0 | \underline{y}\right)$,

(2) specifying k is the same as specifying the posterior probability of $\boldsymbol{\pi_A} = k/(1 + k)$ at which A is accepted.

(3) in the one sample case with τ known and $\pi(\mu) \propto 1$, $W\left(\underline{\boldsymbol{y}}, \mu_0\right) =$ Bayes factor because then $\pi\left(\mu|\underline{\boldsymbol{y}}\right) = f\left(\underline{\boldsymbol{y}}|\mu\right)$ numerically.

& Device

U.S. FOOD & DRUG ADMINISTRATION

Questions?

Gene.Pennello@fda.hhs.gov