Section 12: Hypothesis testing

ARE 210

November 21, 2017

The section notes are available on the section Github at github.com/johnloeser/are210 in the "section12" folder.

1 Definitions

- Hypothesis testing setup
  - $X \sim P \in \mathbf{P}$
  - Two hypotheses: H: $P \in \mathbf{P}_H$ or K: $P \in \mathbf{P}_K$ (where $\mathbf{P}_H \cap \mathbf{P}_K = \emptyset$)
    * Typically say H = null, K = alternative
    * Typically parametrize $\mathbf{P}_H = \{P_\theta : \theta \in \Theta_H\}$, $\mathbf{P}_K = \{P_\theta : \theta \in \Theta_K\}$
  - **Test** $\delta : X \to [0, 1]$
    * δ takes data, says fail to reject null ($\delta(X) = 0$) or reject null ($\delta(X) = 1$)
    * Allow δ to represent more general tests, e.g. $\delta(X) = \frac{1}{2} \ \forall \ X$ represents flipping a coin
  - **Power function** $B_\delta(\theta) \equiv \mathbf{E}_\theta[\delta(X)]$
  - **Size** $\sup_{\theta \in \Theta_H} B_\delta(\theta)$
  - A test is **level-**α if its size is at most α
  - A test ϕ^* is **uniformly most powerful level-**α if 1) ϕ^* is level-α, and 2) for all level α tests ϕ and all $\theta_K \in \Theta_K$, $B_{\phi^*}(\theta_K) \geq B_\phi(\theta_K)$.
  - For a family of tests δ_α, suppose there exists T, t_α such that $\delta_\alpha(X) = \mathbf{1}\{T(X) \geq t_\alpha\}$. Then the **p-value** is $\inf\{\alpha : \delta_\alpha(X) = 1\}$.
  - **Neyman-Pearson Lemma** Suppose $\Theta_H \equiv \{\theta_0\}$ and $\Theta_K \equiv \{\theta_1\}$ are singletons, and $p(x; \theta_i)$ is the pdf or pmf. Then define $\phi_{k,\epsilon}(x) = \mathbf{1}\{p(x, \theta_1)/p(x, \theta_0) > k\} + \epsilon \mathbf{1}\{p(x, \theta_1)/p(x, \theta_0) = k\}$ for some $\epsilon \in \{0, 1\}$. Then 1) $\phi_{k,\epsilon}$ is the uniformly most powerful level-$\mathbf{E}_{\theta_0}[\phi_{k,\epsilon}(x)]$ test, 2) $\forall \ \alpha \in [0, 1]$, $\exists$ a most powerful level-α test of the form $\phi_{k,\epsilon}$, and 3) if ϕ is the most powerful level-α test, then $\exists \ k, \epsilon$ such that $\phi(\cdot) = \phi_{k,\epsilon}(\cdot)$.
  - $\{P_\theta : \theta \in \Theta\}$, where $\Theta \subset \mathbf{R}$, is an **increasing monotone likelihood ratio family** in $T(X)$ if $\frac{p(x,\theta_2)}{p(x,\theta_1)}$ is increasing in $T(X)$ for all (θ_1, θ_2)

* $*$ For exponential family $p(x,\theta) = h(x)\exp(\eta(\theta)T(x) - \beta(\theta))$, η is increasing $\Rightarrow p(\cdot,\theta)$ is increasing monotone likelihood ratio in $\frac{1}{n}\sum_{i=1}^n T(X_i)$
* $*$ $\delta_t(X) = \mathbf{1}\{T(X) > t\}$ is the uniformly most powerful test of H : $\theta_H = \{\theta : \theta \leq \theta_0\}$ against K : $\theta_K = \{\theta : \theta > \theta_0\}$ of level-$\mathbf{E}_{\theta_0}[\delta_t(X)]$

- δ is **unbiased** if $B_\phi(\theta) \leq \alpha \ \forall \ \theta \in \Theta_H$ and $B_\phi(\theta) \geq \alpha \ \forall \ \theta \in \Theta_K$
- δ is **uniformly most powerful unbiased level-**α if 1) δ is unbiased and level-α, and 2) $\forall \ \theta \in \Theta_K$, $B_\delta(\theta) \geq B_\phi(\theta)$ $\forall$ unbiased level-α ϕ
- δ is α**-similar** on $\Theta^* \subset \Theta$ if $B_\delta(\theta) = \alpha \ \forall \ \theta \in \Theta^*$
- Suppose δ is unbiased level-α, and $B_\delta(\theta)$ is continuous $\forall \ \theta \in \text{bdry}(\Theta)$. Then 1) $B_\delta(\theta) = \alpha \ \forall \ \theta \in \text{bdry}(\Theta)$, and 2) if δ is uniformly most powerful among all α-similar tests, then ϕ^* is uniformly most powerful level-α.

2 Practice questions

1. Assume $\{X_i\}_{i=1}^n$ iid $P \in \mathbf{P} = \{P : V_P[X_i] = \sigma^2\}$. Let $\mathbf{P}_H = \{P : V_P[X_i] = \sigma^2, \mathbf{E}_P[X_i] = \mu\}$, representing the null hypothesis that the mean of X_i is μ.
   (a) Calculate $V[\overline{X}]$, where $\overline{X} = \sum_{i=1}^n X_i/n$.
   (b) Assume $P \in \mathbf{P}_H$. Place a lower bound on $P\left[\left|(\overline{X}-\mu)/\sqrt{V[\overline{X}]}\right|\right]$.
   (c) Let $\delta_\alpha = \mathbf{1}\left\{\left|(\overline{X}-\mu)/\sqrt{V[\overline{X}]}\right| > t_{1-\alpha/2}\right\}$. Derive the maximum threshold $t_{1-\alpha/2}$ such that δ_α is size α. Compare this threshold to $Z_{1-\alpha/2}$.[1]
   (d) δ_α is not unbiased, provide a counter example. What intuition does this give you about the set of possible unbiased tests in this context? Think about what bdry($\mathbf{P}_H$) looks like.
   (e) Construct a useful bound on $\inf_{P\in\mathbf{P}_{\mu+k}}\{\mathbf{E}_P[\delta_\alpha]\}$, where $\mathbf{P}_{\mu+k} \equiv \{P : \mathbf{E}_P[X_i] = \mu + k\}$.

[1] Recall $Z_{.84} \approx 1$, $Z_{.97} \approx 2$, $Z_{.998} \approx 3$.