Achievable Error Probabilities for Composite Hypothesis Testing

Pierre Moulin
University of Illinois
Urbana, IL

Yen-Wei Huang
Microsoft
Redmond, WA

CUHK
November 17, 2014

Outline

I. Introduction

II. Likelihood Ratio Threshold Tests

III. Generalized NP Tests

IV. Asymptotics

I. Introduction

Simple Hypothesis Testing

- Observe length-n sequence $\mathbf{y} \in \mathcal{Y}^n$, test simple hypotheses

$$H_0 : \; \mathbf{Y} \sim \mathbb{P} \quad vs. \quad H_1 : \; \mathbf{Y} \sim \mathbb{Q}$$

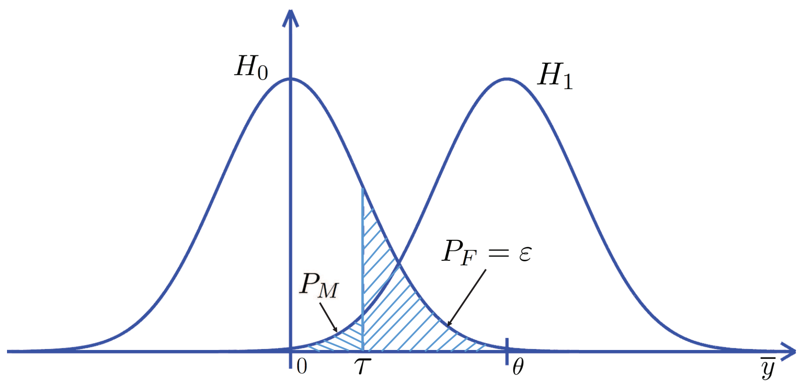
- Randomized decision rule $\delta(\mathbf{y}) \triangleq \Pr\{\text{Say } H_0 | \mathbf{Y} = \mathbf{y}\}, \mathbf{y} \in \mathcal{Y}^n$
- Neyman-Pearson test: minimize false-alarm error probability $\mathbb{E}_Q[\delta(\mathbf{Y})]$ subject to constraint $\mathbb{E}_P[\delta(\mathbf{Y})] \leq 1 - \epsilon$ on miss prob.
- The value of the minimum is denoted by $\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q})$
- Randomized Likelihood Ratio Tests are Neyman-Pearson tests

$$\delta(\mathbf{y}) = \Pr\{\text{Say } H_0 | \mathbf{Y} = \mathbf{y}\} = \begin{cases} 1 & : \; L(\mathbf{y}) > \eta \\ \gamma & : \; L(\mathbf{y}) = \eta \\ 0 & : \; L(\mathbf{y}) < \eta \end{cases}$$

with likelihood ratio $L(\mathbf{Y}) = \frac{d\mathbb{P}(\mathbf{Y})}{d\mathbb{Q}(\mathbf{Y})}$ and threshold η

Gaussian Hypothesis Testing

- Here $\mathcal{Y} = \mathbb{R}$, $\mathbb{P} = P^n$, $\mathbb{Q} = Q^n$ with $P = \mathcal{N}(0,1)$, $Q = \mathcal{N}(\theta, 1)$



- Sample mean $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \Rightarrow \mathcal{N}(0, 1/n)$ vs $\mathcal{N}(\theta, 1/n)$
- Assume $\theta > 0$
- NP test $\delta_{\text{NP}}(\overline{Y}) = \mathbb{1}\{\overline{Y} \le \tau\}$, threshold $\tau = n^{-1/2}\mathcal{Q}^{-1}(\epsilon)$

- $\beta_{1-\epsilon}(\mathbb{P},\mathbb{Q}) = \mathbb{Q}\{\overline{Y} \le \tau\} = \mathcal{Q}(\theta\sqrt{n} - \mathcal{Q}^{-1}(\epsilon))$
- Asymptotics: use $\mathcal{Q}(t) \sim \frac{\exp\{-t^2/2\}}{t\sqrt{2\pi}}$ as $t \to \infty$, then

$$\beta_{1-\epsilon}(\mathbb{P},\mathbb{Q}) \sim \frac{\exp\{-\frac{n}{2}\theta^2 + \theta\sqrt{n}\mathcal{Q}^{-1}(\epsilon) - [\mathcal{Q}^{-1}(\epsilon)]^2/2\}}{\theta\sqrt{2\pi n}}$$

- Error exponent $= \frac{1}{2}\theta^2 = D(P\|Q)$
  (as expected from Stein's lemma)

Composite Hypothesis Testing

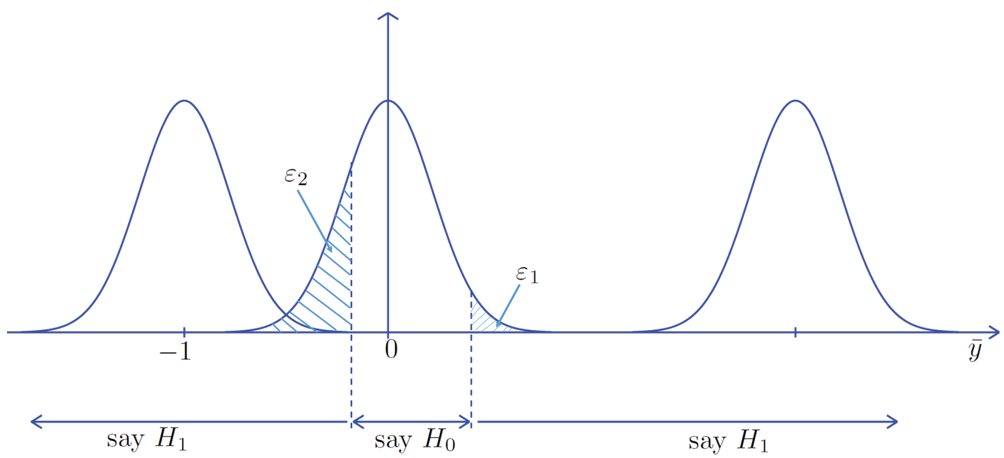
- Observe length-n sequence $\mathbf{y} \in \mathcal{Y}^n$, test simple hypothesis H_0 against composite hypothesis H_1 with k alternatives:

$$\begin{aligned} H_0 : &\quad \mathbf{Y} \sim \mathbb{P} \\ H_1 : &\quad \mathbf{Y} \sim \mathbb{Q}_j \text{ for some } 1 \leq j \leq k \end{aligned}$$

- Applications to statistics, outlier hypothesis testing, and multiuser information theory.
- This work does not address problems where the number of alternatives is uncountable.
- Will assume $\mathbb{P} = \prod_{i=1}^n P_i$ and $\mathbb{Q}_j = \prod_{i=1}^n Q_{ji}$

Example: Gaussian hypothesis testing, $k = 2$

- Here $\mathcal{Y} = \mathbb{R}$, $\mathbb{P} = P^n$, and $\mathbb{Q}_j = Q_j^n$ for $j = 1, 2$ with $P = \mathcal{N}(0, 1)$, $Q_1 = \mathcal{N}(\theta_1, 1)$, $Q_2 = \mathcal{N}(-1, 1)$.
- Consider tests that are functions of $\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$ and have power $1 - \epsilon$



- $\epsilon_1 + \epsilon_2 = \epsilon \quad \Rightarrow$ free parameter ϵ_1
- What is an optimal test at significance level $1 - \epsilon$?

- Compare with oracle LRT which "knows j"
- Asymptotics of oracle LRT:

$$\beta_{1-\epsilon}(\mathbb{P},\mathbb{Q}_1) \sim \frac{\exp\{-\frac{n}{2}\theta_1^2 + \theta_1\sqrt{n}\mathcal{Q}^{-1}(\epsilon) - [\mathcal{Q}^{-1}(\epsilon)]^2/2\}}{\theta_1\sqrt{2\pi n}}$$

$$\beta_{1-\epsilon}(\mathbb{P},\mathbb{Q}_2) \sim \frac{\exp\{-\frac{n}{2} + \sqrt{n}\mathcal{Q}^{-1}(\epsilon) - [\mathcal{Q}^{-1}(\epsilon)]^2/2\}}{\sqrt{2\pi n}}$$

GLRT

- Generalized Likelihood Ratio $L_{\mathrm{G}}(\mathbf{y}) = \frac{d\mathbb{P}(\mathbf{y})}{\max_{1 \le j \le k} d\mathbb{Q}_j(\mathbf{y})}$
- (Deterministic) GLRT with threshold η:

$$\delta_{\mathrm{GLRT}}(\mathbf{y}) = \begin{cases} 1 & : \; L_{\mathrm{G}}(\mathbf{y}) \ge \eta \\ 0 & : \; L_{\mathrm{G}}(\mathbf{y}) < \eta \end{cases}$$

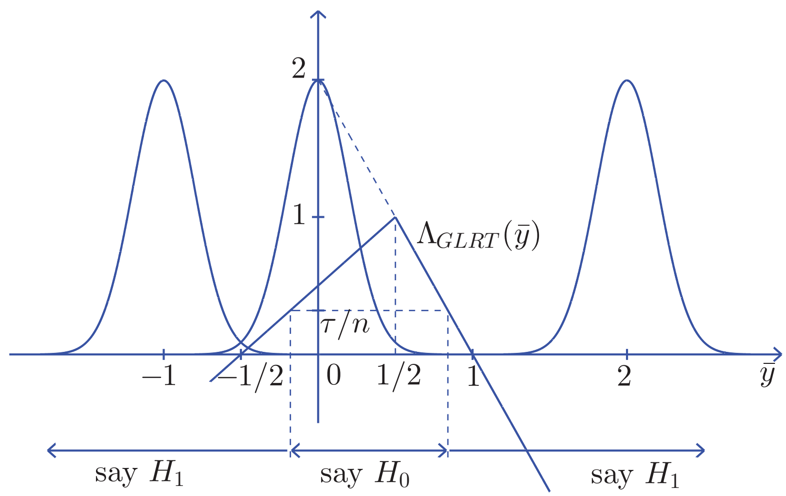
- Equivalently,

$$\delta_{\mathrm{GLRT}}(\mathbf{y}) = \mathbb{1}\left\{ \min_{1 \le j \le k} \frac{d\mathbb{P}(\mathbf{y})}{d\mathbb{Q}_j(\mathbf{y})} \ge \eta \right\}$$

- Widely used, simple to implement, same error exponents as oracle LRT *in some settings*

GLRT in Gaussian setting

- GLRT test statistic $\Lambda_{\mathrm{GLRT}}(\overline{Y}) = \mathbb{1}\left\{\min_{j=1,2} \log \frac{p(\overline{Y})}{q_j(\overline{Y})} \geq \tau_\epsilon\right\}$
$= \mathbb{1}\{-\frac{1}{\sqrt{n}}\mathcal{Q}^{-1}(\epsilon_1) \leq \overline{Y} \leq \frac{\theta_1^2-1}{2\theta_1} + \frac{1}{\theta_1\sqrt{n}}\mathcal{Q}^{-1}(\epsilon_1)\}$
- Nonsymmetric case: assume $\theta_1 > 1$



- Asymptotics of false-positive error probabilities:

$$\begin{aligned} e(1) &= \exp\left\{-\frac{(\theta_1^2+1)^2}{8\theta_1^2}n + O(\sqrt{n})\right\} \\ e(2) &\sim \beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_2) \end{aligned}$$

- $e(2)$ is asymptotically same as for oracle LRT, but error exponent $\frac{(\theta_1^2+1)^2}{8\theta_1^2}$ for $e(1)$ is worse than that for oracle LRT ($\frac{\theta_1^2}{2}$)
- Can we do better?

II. Likelihood Ratio Threshold Tests

Loglikelihood Ratio Vector

- n component loglikelihood-ratio vectors

$$\mathbf{L}_i \triangleq \begin{bmatrix} \log dP_i/dQ_i^1(Y_i) \\ \vdots \\ \log dP_i/dQ_i^k(Y_i) \end{bmatrix}, \quad 1 \le i \le n$$

- loglikelihood-ratio vector

$$\mathbf{Z}_n(\mathbf{Y}) = \sum_{i=1}^{n} \mathbf{L}_i = \begin{bmatrix} \log d\mathbb{P}/d\mathbb{Q}_1(\mathbf{Y}) \\ \vdots \\ \log d\mathbb{P}/d\mathbb{Q}_k(\mathbf{Y}) \end{bmatrix}$$

- mean vector $\mathbf{D}_i = \mathbb{E}_{P_i}[\mathbf{L}_i] = \{D(P_i \| Q_{ji})\}_{j=1}^{k}$
- covariance matrix $\mathsf{V}_i = \mathrm{Cov}_{P_i}(\mathbf{L}_i), \quad i \ge 1$

Likelihood Ratio Threshold Test (LRTT)

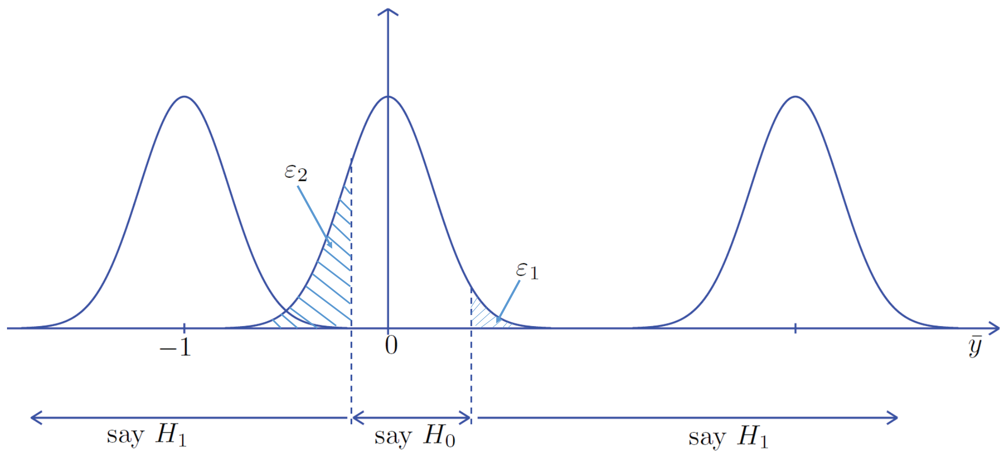
- deterministic test with threshold vector $\boldsymbol{\tau} = [\tau_1, \tau_2, \cdots, \tau_k]'$:

$$\delta_{\text{LRTT}}(\mathbf{y}) \triangleq \mathbb{1}\{\mathbf{Z}_n(\mathbf{y}) \geq \boldsymbol{\tau}\}$$

- GLRT is a special case with $\boldsymbol{\tau} = [\tau, \tau, \cdots, \tau]'$:

$$\delta_{\text{GLRT}}(\mathbf{y}) = \mathbb{1}\left\{\min_{1\leq j\leq k} \frac{d\mathbb{P}(\mathbf{y})}{d\mathbb{Q}_j(\mathbf{y})} \geq \eta\right\}, \quad \eta = e^{\tau}$$

LRTT for Composite Gaussian HT, $k = 2$



- Loglikelihood ratio vector $\mathbf{Z}_n$ has two components

$$Z_{n1} = n\left(-\theta_1\overline{Y} + \theta_1^2/2\right), \quad Z_{n2} = n\left(\overline{Y} + 1/2\right),$$

- LRTT: choose ϵ_1 and ϵ_2 s.t. $\epsilon_1 + \epsilon_2 = \epsilon$ and thresholds

$$\tau_1 = \frac{n}{2}\theta_1^2 - \theta_1\sqrt{n}\mathcal{Q}^{-1}(\epsilon_1), \quad \tau_2 = \frac{n}{2} - \sqrt{n}\mathcal{Q}^{-1}(\epsilon_2),$$

then $\delta_{\mathrm{LRTT}}(\overline{Y}) = \mathbb{1}\{-\mathcal{Q}^{-1}(\epsilon_2) \le \sqrt{n}\,\overline{Y} \le \mathcal{Q}^{-1}(\epsilon_1)\}$

- Recall asymptotics of oracle LRT:

$$\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_1) \quad \sim \quad \frac{\exp\{-\frac{n}{2}\theta_1^2 + \theta_1\sqrt{n}\mathcal{Q}^{-1}(\epsilon) - [\mathcal{Q}^{-1}(\epsilon)]^2/2\}}{\theta_1\sqrt{2\pi n}}$$

$$\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_2) \quad \sim \quad \frac{\exp\{-\frac{n}{2} + \sqrt{n}\mathcal{Q}^{-1}(\epsilon) - [\mathcal{Q}^{-1}(\epsilon)]^2/2\}}{\sqrt{2\pi n}}.$$

- LRTT power $\mathbb{E}_P[\delta_{\mathrm{LRTT}}(\overline{Y})] = 1 - \epsilon$ and false-positive probs

$$e(1) \sim \beta_{1-\epsilon_1}(\mathbb{P}, \mathbb{Q}_1), \quad e(2) \sim \beta_{1-\epsilon_2}(\mathbb{P}, \mathbb{Q}_2)$$

- Same error exponents as "oracle" LRT but the 2nd-order terms are worse since $\epsilon_1, \epsilon_2 < \epsilon \quad \Rightarrow \mathcal{Q}^{-1}(\epsilon_1), \mathcal{Q}^{-1}(\epsilon_2) > \mathcal{Q}^{-1}(\epsilon)$

- LRTT outperforms GLRT here

III. Generalized NP Tests

Error Probabilities for Composite HT

- Randomized decision rule $\delta(\mathbf{y}) \triangleq \Pr\{\text{Say } H_0 | \mathbf{Y} = \mathbf{y}\}, \mathbf{y} \in \mathcal{Y}^n$
- Constrain test power $\mathbb{E}_{\mathbb{P}}[\delta(\mathbf{Y})] \geq 1 - \epsilon$ for some fixed $\epsilon \in (0, 1)$
- False-positive error probabilities $e_j = \mathbb{E}_{\mathbb{Q}_j}\left[\delta(\mathbf{Y})\right], \, 1 \leq j \leq k$
- Set of achievable $\{e_j\}_{j=1}^k$ for power $(1-\epsilon)$ tests:

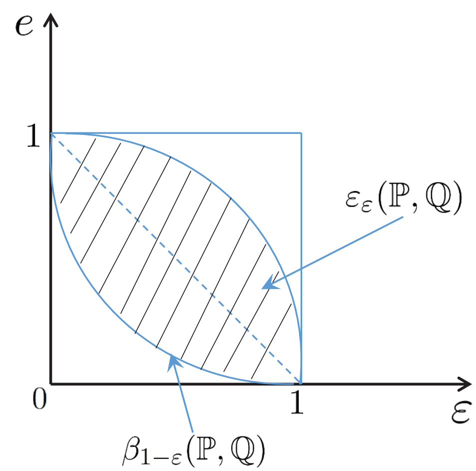
$$\mathcal{E}_\epsilon\left(\mathbb{P}, \{\mathbb{Q}_j\}_{j=1}^k\right) \triangleq \Big\{[e_1, \ldots, e_k]' : \exists \text{test } \delta \text{ s.t.}$$
$$\mathbb{E}_{\mathbb{P}}\left[\delta(\mathbf{Y})\right] \geq 1 - \epsilon \text{ and } \mathbb{E}_{\mathbb{Q}_j}\left[\delta(\mathbf{Y})\right] = e_j, 1 \leq j \leq k\Big\}.$$

- The point $\{\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_j)\}_{j=1}^k$ is achievable only in the rare problems where a *Uniformly Most Powerful* test exists.

- If $k = 1$, then (Neyman-Pearson)

$$\mathcal{E}_\epsilon(\mathbb{P}, \mathbb{Q}_1) = [\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_1), 1 - \beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}_1)]$$

  Optimal (non-dominated) test is randomized likelihood ratio test (LRT).



- For $k \geq 2$, nondominated tests are of interest. They are "**generalized NP tests**" (Lehmann) but generally **not** LRTs.

$$\Rightarrow \mathcal{E}_\epsilon^{\text{GNP}}\left(\mathbb{P}, \{\mathbb{Q}_j\}_{j=1}^k\right) \subset \mathcal{E}_\epsilon\left(\mathbb{P}, \{\mathbb{Q}_j\}_{j=1}^k\right)$$

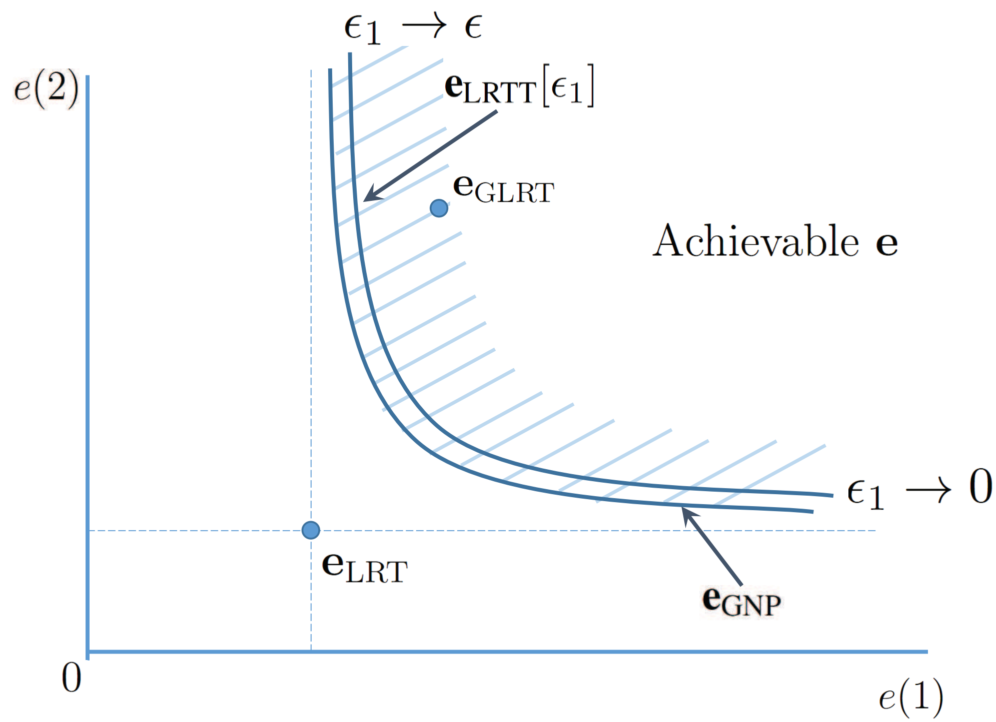
- Problem here: derive precise asymptotics of nondominated set $\mathcal{E}_\epsilon^{\mathrm{GNP}}\left(\mathbb{P},\{\mathbb{Q}_j\}_{j=1}^k\right)$ when $\mathbb{P}=\prod_{i=1}^n P_i$ and $\mathbb{Q}_j=\prod_{i=1}^n Q_{ji}$.
- Will do so by relating GNP tests to LRTTs
- Achievable error vectors for power $(1-\epsilon)$ LRTTs:

$$\mathcal{E}_\epsilon^{\mathrm{LRTT}}\left(\mathbb{P},\{\mathbb{Q}_j\}_{j=1}^k\right) \triangleq \Big\{[e_1,\ldots,e_k]' : \exists \text{LRTT } \delta_{\mathrm{LRTT}} \text{ s.t.}$$
$$\mathbb{E}_{\mathbb{P}}\left[\delta_{\mathrm{LRTT}}(\mathbf{Y})\right] \geq 1-\epsilon \text{ and } \mathbb{E}_{\mathbb{Q}_j}\left[\delta_{\mathrm{LRTT}}(\mathbf{Y})\right] = e_j, 1 \leq j \leq k\Big\}$$

- Clearly

$$\mathcal{E}_\epsilon^{\mathrm{LRTT}}\left(\mathbb{P},\{\mathbb{Q}_j\}_{j=1}^k\right) \subset \mathcal{E}_\epsilon\left(\mathbb{P},\{\mathbb{Q}_j\}_{j=1}^k\right)$$

- Achievable false-positive error probabilities for $k = 2$:



Characterization of GNP Tests

- **Proposition 1**. (Variation on Lehmann's Theorem 3.6.1). The set $\mathcal{E}$ of achievable error probabilities

$$\Big[\mathbb{E}_{\mathbb{Q}_1}[\delta(\mathbf{Y})], \quad \cdots, \quad \mathbb{E}_{\mathbb{Q}_k}[\delta(\mathbf{Y})], \quad 1 - \mathbb{E}_{\mathbb{P}}[\delta(\mathbf{Y})]\Big]' \in [0,1]^{k+1}$$

for some test δ is convex and closed. If $[e_1^{\mathrm{FP}}, \ldots, e_k^{\mathrm{FP}}, e^{\mathrm{FN}}]'$ is a minimal point in $\mathcal{E}$ with $e^{\mathrm{FN}} \in (0,1)$, then $\exists$ a nonzero $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_k]' \geq 0$ and a gen'd NP test δ_{GNP} satisfying

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\delta_{\mathrm{GNP}}(\mathbf{Y})] &= 1 - e^{\mathrm{FN}}, \\ \mathbb{E}_{\mathbb{Q}_j}[\delta_{\mathrm{GNP}}(\mathbf{Y})] &= e_j^{\mathrm{FP}}, \quad j = 1, \ldots, k, \end{aligned}$$

and

$$\begin{aligned} \delta_{\mathrm{GNP}}(\mathbf{y}) = 1 \quad &\text{when} \quad d\mathbb{P}(\mathbf{y}) > \textstyle\sum_{j=1}^k \alpha_j d\mathbb{Q}_j(\mathbf{y}) \\ \delta_{\mathrm{GNP}}(\mathbf{y}) = 0 \quad &\text{when} \quad d\mathbb{P}(\mathbf{y}) < \textstyle\sum_{j=1}^k \alpha_j d\mathbb{Q}_j(\mathbf{y}). \end{aligned}$$

Relation to LRTTs

- By Prop. 1, any δ_{GNP} is parameterized by nonnegative $\{\alpha_j\}_{j=1}^k$.
- Fix arbitrarily small $\eta > 0$. Consider two LRTTs $\delta_{\text{LRTT}}^{\text{in}}$ and $\delta_{\text{LRTT}}^{\text{out}}$ with threshold vectors

$$\tau_j^{\text{in}} = \ln\left[(k+\eta)\alpha_j\right], \quad \tau_j^{\text{out}} = \ln \alpha_j, \quad 1 \le j \le k.$$

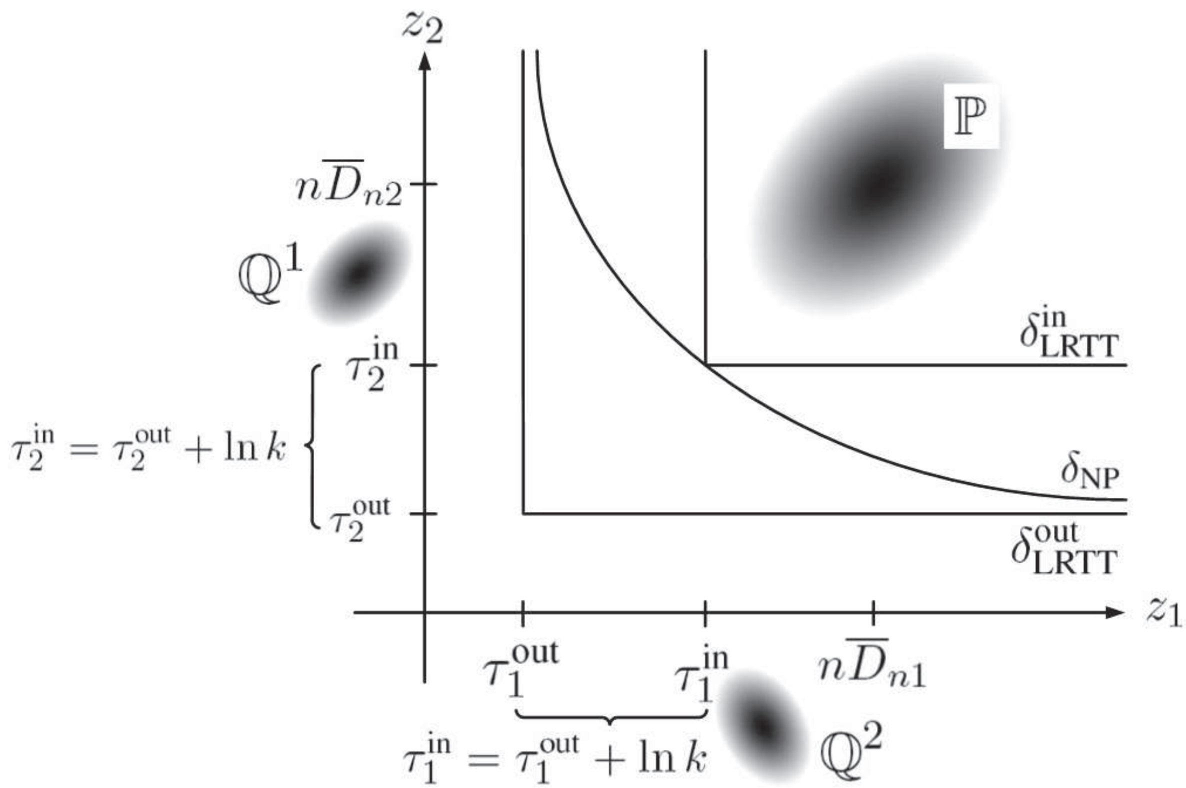
- **Proposition 2**. The power and the false-positive error vector of δ_{GNP} can be sandwiched as follows:

$$\mathsf{E}_{\mathbb{P}}[\delta_{\text{LRTT}}^{\text{in}}(\mathbf{Y})] \le \mathsf{E}_{\mathbb{P}}[\delta_{\text{GNP}}(\mathbf{Y})] \le \mathsf{E}_{\mathbb{P}}[\delta_{\text{LRTT}}^{\text{out}}(\mathbf{Y})]$$

and

$$\mathbf{e}(\delta_{\text{LRTT}}^{\text{in}}) \le \mathbf{e}(\delta_{\text{GNP}}) \le \mathbf{e}(\delta_{\text{LRTT}}^{\text{out}}).$$

- Distributions of $\mathbf{Z}_n$ when $k = 2$:



$(k = 2)$

IV. Asymptotics

Related LD Work

- $k = 1$: Moulin (2013):

$$\beta_{1-\epsilon}(\mathbb{P}, \mathbb{Q}) = e^{-\sum_{i=1}^{n} D(P_i \| Q_i) + \sqrt{\sum_{i=1}^{n} V(P_i \| Q_i)}\, \mathcal{Q}^{-1}(\epsilon) - \frac{1}{2}\log n + c + o(1)}$$

  refining Theorem 3.1 by Strassen (1962) for deterministic tests

- Borovkov and Rogozin (1965), Iltis (1995), Petrovskii (1996), Chaganty and Sethuraman (1996) derived multidimensional strong large deviations theorems

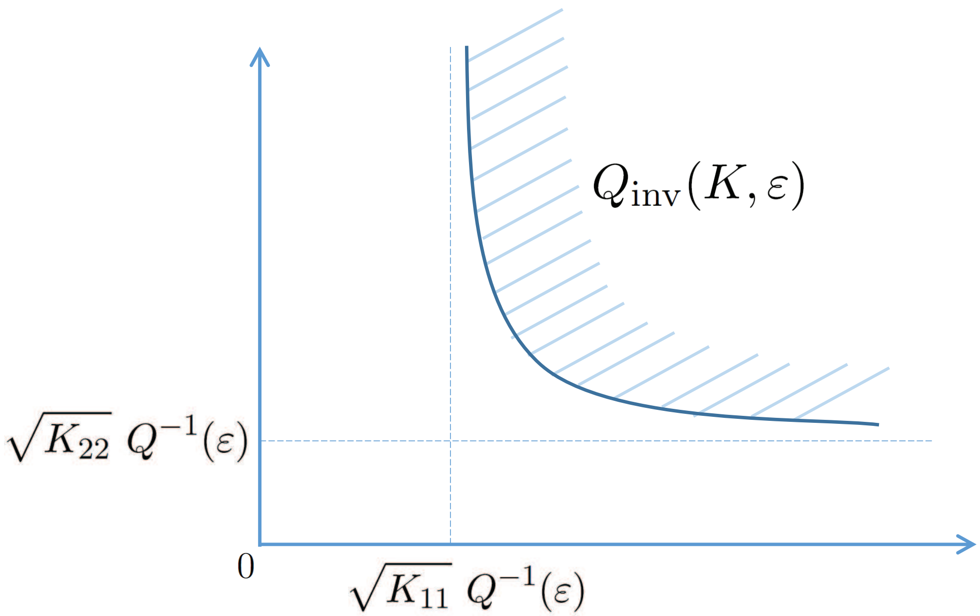
$$\mathbb{P}\left\{\sum_{i=1}^{n} \mathbf{Z}_i \in \mathcal{D}_n\right\} = \exp\{\cdots\}$$

  for rare events, where $\mathbf{Z}_i$, $i \geq 1$ are independent $\mathbb{R}^k$ valued random variables, and $\mathcal{D}_n$ is a sequence of subsets of $\mathbb{R}^k$.

The Set $\mathcal{Q}_{\text{inv}}$

- Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathsf{K})$ and $\epsilon \in (0,1)$. Then

$$\mathcal{Q}_{\text{inv}}(\mathsf{K}, \epsilon) \triangleq \left\{ \boldsymbol{\tau} \in \mathbb{R}^k : \Pr(\mathbf{X} \leq \boldsymbol{\tau}) \geq 1 - \epsilon \right\}$$



Assumptions

(A1) (Bounded moments). There exists $\eta > 0$ s.t. $\forall n \geq 1$,

$$\begin{aligned}
\eta\,\mathbf{1} < \overline{\mathbf{D}}_n &\triangleq \frac{1}{n}\sum_{i=1}^{n}\mathbf{D}_i < \frac{1}{\eta}\,\mathbf{1}, \\
\eta\,\mathsf{I}_k < \overline{\mathsf{V}}_n &\triangleq \frac{1}{n}\sum_{i=1}^{n}\mathsf{V}_i < \frac{1}{\eta}\,\mathsf{I}_k, \\
\overline{T}_n &\triangleq \frac{1}{n}\mathbb{E}_{\mathbb{P}}[\|\mathbf{Z}_n - n\overline{\mathbf{D}}_n\|_2^3] < \frac{1}{\eta}.
\end{aligned}$$

(A2) $\mathbf{L}_i, i \geq 1$ are strongly nonlattice random vectors.

LRTTs

Theorem 1: Let assumptions **(A1)** and **(A2)** hold. Then

$$\mathbf{e} \in \mathcal{E}_{\epsilon}^{\mathrm{LRTT}}\left(\mathbb{P}, \{\mathbb{Q}_j\}_{j=1}^{k}\right) \quad \Leftrightarrow \; \mathbf{e} = \exp\left\{-n\overline{\mathbf{D}}_n + \sqrt{n}\mathbf{b} - \frac{1}{2}\log n\,\mathbf{1} + \mathbf{c} + o_{\mathbf{b},\mathbf{c}}(1)\right\}$$

where $\mathbf{b} \in \partial\mathcal{Q}_{\mathrm{inv}}(\overline{\mathsf{V}}_n, \epsilon)$ and $\mathbf{c}$ satisfies a linear inequality constraint

Sketch of the Proof.

- Let $\mathbf{Z}_n^* \sim \mathcal{N}(n\overline{\mathbf{D}}_n, n\overline{\mathsf{V}}_n)$.
  By the multidimensional Berry-Esséen theorem,

$$\forall \boldsymbol{\tau}_n : \quad |\mathbb{P}\{\mathbf{Z}_n \geq \boldsymbol{\tau}_n\} - \Pr\{\mathbf{Z}_n^* \geq \boldsymbol{\tau}_n\}| \leq \gamma_n = O(n^{-1/2})$$

- Consider **any** LRTT with threshold vector $\boldsymbol{\tau}_n$ and $\mathbb{P}\{\mathbf{Z}_n \geq \boldsymbol{\tau}_n\} \geq 1-\epsilon$. Then $\Pr\{\mathbf{Z}_n^* \geq \boldsymbol{\tau}_n\} \geq 1-\epsilon-\gamma_n$, hence

$$\boldsymbol{\tau}_n \in n\overline{\mathbf{D}}_n - \sqrt{n}\mathcal{Q}_{\mathrm{inv}}(\overline{\mathsf{V}}_n, \epsilon + \gamma_n).$$

- Define $\mathbf{U}_n = n^{-1/2}(\mathbf{Z}_n - \boldsymbol{\tau}_n)$. For $1 \le j \le k$ we have

$$\begin{aligned}
\mathbb{Q}_j[\mathbf{Z}_n \ge \boldsymbol{\tau}_n] &= \mathbb{E}_{\mathbb{Q}_j}[\mathbb{1}\{\mathbf{Z}_n \ge \boldsymbol{\tau}_n\}] \\
&= \mathbb{E}_{\mathbb{P}}[e^{-Z_{nj}}\mathbb{1}\{\mathbf{Z}_n \ge \boldsymbol{\tau}_n\}] \\
&= e^{-\tau_j}\mathbb{E}_{\mathbb{P}}[e^{-\sqrt{n}\,U_{nj}}\mathbb{1}\{\mathbf{U}_n \ge \mathbf{0}\}] \\
&\overset{(*)}{=} \exp\left\{-\tau_j - \frac{1}{2}\log n + O(1)\right\}
\end{aligned}$$

  where (*) is proven using a variation on Chaganty and Sethuraman (1996).

- Use $\boldsymbol{\tau}_n \in n\overline{\mathbf{D}}_n - \sqrt{n}\mathcal{Q}_{\text{inv}}(\overline{\mathsf{V}}_n, \epsilon + \gamma_n)$ and Taylor expansion to conclude that the error vector

$$\mathbf{e} \in \exp\left\{-n\overline{\mathbf{D}}_n + \sqrt{n}\mathcal{Q}_{\text{inv}}(\overline{\mathsf{V}}_n, \epsilon) - \frac{1}{2}\log n\,\mathbf{1} + O(1)\right\}$$

GNP Tests

- **Theorem 2**: Let assumptions **(A1)** and **(A2)** hold. Then

$$\mathbf{e} \in \mathcal{E}_\epsilon^{\mathrm{LRTT}}\left(\mathbb{P}, \{\mathbb{Q}_j\}_{j=1}^k\right) \quad \Leftrightarrow$$

$$\mathbf{e} \quad = \quad \exp\left\{-n\overline{\mathbf{D}}_n + \sqrt{n}\mathbf{b} - \frac{1}{2}\log n\,\mathbf{1} + \mathbf{c} - \mathbf{d}_n + o_{\mathbf{b},\mathbf{c}}(1)\right\}$$

  where $\mathbf{b} \in \partial\mathcal{Q}_{\mathrm{inv}}(\overline{\mathsf{V}}_n, \epsilon)$; $\mathbf{c}$ satisfies a linear inequality constraint; and $\mathbf{0} \leq \mathbf{d}_n \leq \ln k\mathbf{1}$

- Proof: application of Theorem 1 and Prop. 2

Conclusion

- LRTTs are simple and powerful for composite HT with k alternatives
- While optimal decision rules for composite hypothesis testing are generally not LRTTs, precise asymptotic characterization of achievable errors is possible.
- Achievable error probabilities are asymptotically within $\ln k$ of achievable error probabilities for deterministic LRTTs.