CHAPTER 6.

TESTING

1. NEYMAN - PEARSON TESTS

1.1. Basic Notation. Consider the hypothesis testing problem as in 5.1.4 and 5.5.4, but with $\Theta_0 = \{0\}$, $\Theta_1 = \{1\}$ (simple hypotheses). Let ϕ be a critical function (or decision rule);
let α = size or level $\equiv E_0\phi(X)$;
let β = power = $E_1\phi(X)$.

1.2. Theorem. (Neyman - Pearson lemma). Let P_0 and P_1 have densities p_0 and p_1 wrt some dominating measure μ (recall that $\mu = P_0 + P_1$ always works). Let $0 \le \alpha \le 1$. Then:

(i) There exists a constant k and a critical function ϕ of the form

(1) $$\phi(x) = \begin{cases} 1 & \text{if } p_1(x) > k\, p_0(x) \\ 0 & \text{if } p_1(x) < k\, p_0(x) \end{cases}$$

such that

(2) $$E_0\phi(X) = \alpha .$$

(ii) The test of (1) and (2) is a *most powerful* α level test of P_0 versus P_1 .

(iii) If ϕ is most powerful level α of P_0 versus P_1 , then it must be of the form (1) a.e. μ . It also satisfies (2) unless there is a test of size $< \alpha$ with power $= 1$.

1.3. Corollary. If $0 < \alpha < 1$ and β is the power of the most powerful level α test, then $\alpha < \beta$ unless $P_0 = P_1$.

Proof. Let $0 < \alpha < 1$.
(i) Now

$$P_0(\, p_1(X) > kp_0(X)\,) = P_0(\, Y \equiv p_1(X)/p_0(X) > c\,) = 1 - F_Y(c) .$$

Let $k \equiv \inf\{c : 1 - F_Y(c) < \alpha\}$, and if $P_0(Y = k) > 0$, let $\gamma \equiv (\alpha - P_0(Y > k))/P_0(Y = k)$. Thus with

$$\phi(x) = \begin{cases} 1 & \text{if } p_1(x) > kp_0(x) \\ \gamma & \text{if } p_1(x) = kp_0(x) \\ 0 & \text{if } p_1(x) < kp_0(x) \end{cases} ,$$

we have

$$E_0\phi(X) = P_0(Y > k) + \gamma P_0(Y = k) = \alpha .$$

(ii) Let ϕ^* be another test with $E_0\phi^* \le \alpha$. Now

$$\int_{\mathbf{X}} (\phi - \phi^*)(p_1 - kp_0)\, d\mu$$

$$= \int_{[\phi - \phi^* > 0] \cup [\phi - \phi^* < 0]} (\phi - \phi^*)(p_1 - kp_0)\, d\mu \ \geq \ 0 \, ,$$

and this implies that

$$\begin{aligned} \beta_\phi - \beta_{\phi^*} &= \int_{\mathbf{X}} (\phi - \phi^*)\, p_1 d\mu \\ &\geq k \int_{\mathbf{X}} (\phi - \phi^*)\, p_0 \, d\mu \ = \ k(\alpha - E_0 \phi^*) \ \geq \ 0 \, . \end{aligned}$$

Thus ϕ is most powerful.

(iii) Let ϕ^* be most powerful level α. Define ϕ as in (i). Then

$$\begin{aligned} \int_{\mathbf{X}} (\phi - \phi^*)(p_1 - kp_0)\, d\mu &= \int_{[\phi \neq \phi^*] \cap [p_1 \neq kp_0]} (\phi - \phi^*)(p_1 - kp_0)\, d\mu \\ &= \begin{cases} \geq 0 & \text{as in (ii)} \\ > 0 & \text{if } \ \mu([\phi \neq \phi^*] \cap [p_1 \neq kp_0]) > 0 \end{cases} \\ &= 0 \qquad \text{since } > 0 \end{aligned}$$

contradicts ϕ^* being most powerful. Thus $\mu([\phi \neq \phi^*] \cap [p_1 \neq kp_0]) = 0$. Thus $\phi^* = \phi$ on the set where $p_1 \neq kp_0$. □

Corollary proof. $\phi^\#(x) \equiv \alpha$ has power α, so $\beta \geq \alpha$. If $\beta = \alpha$, then $\phi^\# \equiv \alpha$ is in fact most powerful; and hence (iii) shows that $\phi(x) = \alpha$ satisfies (i); that is, $p_1(x) = kp_0(x)$ a.e. μ. Thus $k = 1$ and $P_1 = P_0$. □

- If $\alpha = 0$, let $k = \infty$ and $\phi(x) = 1$ whenever $p_1(x)/p_0(x) = \infty$; this is $\gamma = 1$.
- If $\alpha = 1$, let $k = 0$ and $\gamma = 1$, so that we reject for all x with $p_1(x) > 0$ or $p_0(x) > 0$. □

1.4. Definition. If the family of densities $\{p_\theta : \theta \in [\theta_0, \theta_1] \subset R\}$ is such that $p_{\theta'}(x)/p_\theta(x)$ is nondecreasing in $T(x)$ for each $\theta < \theta'$, then the family is said to have *monotone likelihood ratio* (MLR).

1.5. Definition. A test ϕ is of *size* α if

$$\sup_{\theta \in \Theta_0} E_\theta \phi(X) = \alpha \, .$$

Let $\mathbf{C}_\alpha \equiv \{\phi : \phi \text{ is of size } \alpha\}$. A test ϕ_0 is *uniformly most powerful of size* α (UMP of size α) if it has size α and

$$E_\theta \phi_0(X) \geq E_\theta \phi(X) \qquad \text{for all } \ \theta \in \Theta_1 \quad \text{and all} \quad \phi \in \mathbf{C}_\alpha \, .$$

1.6. Theorem. (Karlin - Rubin). Suppose that X has density p_θ with MLR in $T(x)$.

(i) Then there exists a UMP level α test of $H: \theta \le \theta_0$ versus $K: \theta > \theta_0$ which is of the form

$$\phi(x) = \begin{cases} 1 & \text{if } T(x) > c \\ \gamma & \text{if } T(x) = c \\ 0 & \text{if } T(x) < c \end{cases}$$

with $E_{\theta_0}\phi(X) = \alpha$.

(ii) $\beta(\theta) = E_\theta \phi(X)$ is increasing in θ for $\beta < 1$.

(iii) For all θ' this same test is the UMP level $\alpha' \equiv \beta(\theta')$ test of $H': \theta \le \theta'$ versus $K': \theta > \theta'$.

(iv) For all $\theta < \theta_0$, the test of (i) minimizes $\beta(\theta)$ among tests satisfying $\alpha = E_{\theta_0}\phi$.

Proof. (i) and (ii). The most powerful level α test of θ_0 versus $\theta_1 > \theta_0$ is the ϕ above, by the Neyman - Pearson lemma, which guarantees the existence of C and γ. Thus ϕ is UMP of θ_0 versus $\theta > \theta_0$. According to the NP lemma (ii), this same test is most powerful of θ' versus θ''; thus (ii) follows from the the NP corollary. Thus ϕ is also level α in the smaller class of tests of H versus K; and hence is UMP there also: note that with $\mathbf{C}_\alpha \equiv \{\phi : \sup_{\theta \le \theta_0} E_\theta \phi = \alpha\}$ and $\mathbf{C}_\alpha^{\theta_0} \equiv \{\phi : E_{\theta_0}\phi \le \alpha\}$, $\mathbf{C}_\alpha \subset \mathbf{C}_\alpha^{\theta_0}$.

(iii) The same argument works.

(iv) To minimize power, just apply the NP lemma with inequalities reversed. □

1.6. Examples.

Example 1. (Hypergeometric). Suppose that we sample without replacement n items from a population of N items of which $\theta = D$ are defective. Let $X \equiv$ number of defective items in the sample. Then

$$P_D(X = x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}, \quad \text{for } x = 0 \bigvee (n - N + D), \cdots, D \bigwedge n.$$

Since

$$\frac{p_{D+1}(x)}{p_D(x)} = \frac{D+1}{N-D} \frac{N - D - n + x}{D + 1 - x}$$

is increasing in x, there is MLR in $T(X) = X$. Thus the UMP test of $H: D \le D_0$ versus $K: D > D_0$ rejects H if X is “too big”: $\phi(X) = 1_{[X > c]} + \gamma 1_{[X = c]}$ where

$$P_{D_0}(X > c) + \gamma P_{D_0}(X = c) = \alpha .$$

Example 2. (One - parameter exponential families). Suppose that

$$p_\theta(x) = c(\theta)\exp(Q(\theta)T(x))\, h(x)$$

where $Q(\theta)$ is increasing in θ. Then

$$\phi(X) = \begin{cases} 1 & \text{if } T(X) > c \\ \gamma & \text{if } T(X) = c \\ 0 & \text{if } T(X) < c \end{cases}$$

with $E_{\theta_0}\phi(X) = \alpha$ is UMP level α for testing $H : \theta \le \theta_0$ versus $K : \theta > \theta_0$. [See pages 70 - 71 in TSH for binomial, negative binomial, Poisson, and exponential examples.]

Example 3. (Noncentral t, χ^2, and F distributions). The noncentral t, χ^2, and F distributions have MLR in their noncentrality parameters. See Lehmann, page 223 for the t distribution; see Lehmann problem 7.4, page 312, for the χ^2 and F distributions.

Example 4. The Cauchy location family $p_\theta(x) = \pi^{-1}(1 + (x - \theta)^2)^{-1}$ does *not* have MLR.

1.7. Theorem. (The generalized N - P lemma). Let $f_0, f_1, \cdots, f_m$ be real - valued, μ – integrable functions defined on a Euclidean space $\mathbf{X}$. Let ϕ_0 be any function of the form

$$\phi_0(x) = \begin{cases} 1 & \text{if } f_0(x) > k_1 f_1(x) + \cdots + k_m f_m(x) \\ \gamma(x) & \text{if } f_0(x) = k_1 f_1(x) + \cdots + k_m f_m(x) \\ 0 & \text{if } f_0(x) < k_1 f_1(x) + \cdots + k_m f_m(x) \end{cases}$$

where $0 \le \gamma(x) \le 1$. Then ϕ_0 maximizes

$$\int \phi\, f_0\, d\mu$$

over all ϕ, $0 \le \phi \le 1$ such that

$$\int \phi f_i\, d\mu = \int \phi_0 f_i\, d\mu , \quad i = 1, \cdots, m .$$

If $k_j \ge 0$ for $j = 1, \cdots, m$, then ϕ_0 maximizes $\int \phi f_0\, d\mu$ out of all functions ϕ, $0 \le \phi \le 1$ such that

$$\int \phi f_i\, d\mu \le \int \phi_0 f_i\, d\mu , \quad i = 1, \cdots, m .$$

Proof. Note that

$$\int (\phi_0 - \phi)(f_0 - \sum_{j=1}^m k_j f_j)\, d\mu \ge 0$$

since the integrand is ≥ 0 by the definition of ϕ. Hence

$$\int (\phi_0 - \phi) f_0 \, d\mu \geq \sum_{j=1}^{m} k_j \int (\phi_0 - \phi) f_j \, d\mu \geq 0$$

in either of the above cases, and hence

$$\int \phi_0 f_0 \, d\mu \geq \int \phi f_0 \, d\mu .$$

□

This is a "short form" of the generalized N-P lemma; for a "long form" with more details, see Lehmann, TSH, page 83.

1.8. Example. Suppose that $X_1, \cdots, X_n$ are iid from the Cauchy location family

$$p(x;\theta) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2} ,$$

and consider testing $H : \theta = \theta_0$ versus $\theta > \theta_0$. Can we find a test ϕ of size α such $\phi_\phi(\theta)$ has

$$\frac{d}{d\theta} \beta_\phi(\theta_0) = \frac{d}{d_\theta} E_\theta \phi(X)|_{\theta = \theta_0}$$

maximum?

For any test ϕ the power is given by

$$\beta_\phi(\theta) = E_\theta \phi(\underline{X}) = \int \phi(\underline{x}) \, p(\underline{x};\theta) \, d\underline{x} ,$$

so, if the interchange of $d/d\theta$ and $\int$ is justifiable, then

$$\beta'_\phi(\theta) = \int \phi(\underline{x}) \frac{\partial}{\partial\theta} p(\underline{x};\theta) \, d\underline{x}$$

Thus, by the generalized N-P lemma, any test of the form

$$\phi(\underline{x}) = \begin{cases} 1 & \text{if } \frac{\partial}{\partial\theta} p(\underline{x};\theta_0) > kp(\underline{x};\theta_0) \\ \gamma(x) & \text{if } \frac{\partial}{\partial\theta} p(\underline{x};\theta_0) = kp(\underline{x};\theta_0) \\ 0 & \text{if } \frac{\partial}{\partial\theta} p(\underline{x};\theta_0) < kp(\underline{x};\theta_0) \end{cases}$$

maximizes $\beta'_\phi(\theta_0)$ among all ϕ with $E_{\theta_0} \phi(\underline{X}) \leq \alpha$. This test is said to be *locally most powerful* of size α; cf. Ferguson, section 5.5, page 235. But

$$\frac{\partial}{\partial\theta} p(\underline{x};\theta_0) > kp(\underline{x};\theta_0)$$

is equivalent to

$$\frac{\partial}{\partial\theta} p(\underline{x};\theta_0)/p(\underline{x};\theta_0) \;>\; k\,,$$

or

$$\frac{\partial}{\partial\theta}\log p(\underline{x};\theta_0) \;>\; k\,,$$

or

$$S_n(\theta_0) \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \dot{\mathbf{l}}_\theta(X_i;\theta_0) \;>\; k'\,.$$

Hence, for the Cauchy family (with $\theta_0 \equiv 0$ without loss of generality), since

$$\frac{\partial}{\partial\theta}\log p(\underline{x};\theta) \;=\; \frac{2\,(x-\theta)}{1+(x-\theta)^2}\,,$$

the locally most powerful test is given by

$$\phi(\underline{X}) \;=\; \begin{cases} 1 & \text{if } \; n^{-1/2}\sum_{i=1}^{n} \dfrac{2X_i}{1+X_i^2} \;>\; k' \\ 0 & \text{if } \; n^{-1/2}\sum_{i=1}^{n} \dfrac{2X_i}{1+X_i^2} \;<\; k' \end{cases}$$

where k' is such that $E_0\phi(\underline{X}) = \alpha$. Under $\theta = \theta_0 \equiv 0$, with $Y_i \equiv 2X_i/(1+X_i^2)$,

$$E_0 Y_i \;=\; 0\,, \qquad Var_0(Y_i) \;=\; 1/2\,.$$

Hence, by the CLT, k' may be approximated by $2^{-1/2}z_\alpha$ where $P(Z \ge z_\alpha) = \alpha$.

Note that $x/(1+x^2) \to 0$ as $x \to \infty$. Thus, if $\alpha < 1/2$ so that $k' > 0$, the rejection set of ϕ is a bounded set in R^n; and since the probability that $\underline{X} = (X_1, \cdots, X_n)$ is any given bounded set goes to zero as $\theta \to \infty$, $\beta_\phi(\theta) \to 0$ as $\theta \to \infty$.

1.9. Consistency of Neyman - Pearson tests

Let P and Q be probability measures, and suppose that p and q are their densities wrt a common $\sigma-$finite measure μ on $(\mathbf{X}, \mathbf{A})$. Recall that the Hellinger distance $H(P,Q)$ between P and Q is given by

$$\begin{aligned} H^2(P,Q) \;&=\; \frac{1}{2}\int (\sqrt{p} \;-\; \sqrt{q})^2\, d\mu \\ &=\; 1 \;-\; \int \sqrt{p\,q}\, d\mu \;=\; 1 \;-\; \rho(P,Q) \end{aligned}$$

where

$$\rho(P,Q) \;\equiv\; \int \sqrt{p\,q}\, d\mu$$

is the *affinity* between P and Q .

1.10. Proposition.
$H(P,Q) = 0$ iff $p = q$ a.e. μ iff $\rho(P,Q) = 1$.
Furthermore $\rho(P,Q) = 0$ iff $\sqrt{p} \perp \sqrt{q}$ in the Hilbert space $L_2(\mu)$.

1.11. Proposition. Let $X_1, \cdots, X_n$ be iid P or (Q) with joint densities

$$p_n(\underline{x}) \equiv p(\underline{x}) = \prod_{i=1}^n p(x_i) \quad \text{or} \quad q_n(\underline{x}) \equiv q(\underline{x}) = \prod_{i=1}^n q(x_i) .$$

Then

$$\rho(P^n, Q^n) = [\rho(P,Q)]^n \to 0$$

unless $p = q$ a.e. μ .

1.12. Theorem. (Size and power consistency of N-P type tests). For testing p versus q the test

$$\phi_n(\underline{x}) = \begin{cases} 1 & \text{if} \quad q_n(\underline{x}) > k_n p_n(\underline{x}) \\ 0 & \text{if} \quad q_n(\underline{x}) < k_n p_n(\underline{x}) \end{cases}$$

with $0 < a_1 \le k_n \le a_2 < \infty$ for all $n \ge 1$ is size and power consistent: both probabilities of error $\to 0$ as $n \to \infty$.

Proof. First the type I error probabilities:

$$\begin{aligned} E_P \phi_n(\underline{X}) &= \int \phi_n(\underline{x})\, p_n(\underline{x})\, d\mu(\underline{x}) \\ &= \int \phi_n(\underline{x})\, p_n^{1/2}(\underline{x})\, p_n^{1/2}(\underline{x})\, d\mu(\underline{x}) \\ &\le k_n^{-1/2} \int \phi_n(\underline{x})\, p_n^{1/2}(\underline{x})\, q_n^{1/2}(\underline{x})\, d\mu(\underline{x}) \\ &= k_n^{-1/2} \rho(P^n, Q^n) \\ &= k_n^{-1/2} [\rho(P,Q)]^n \\ &\to 0 \quad \text{as} \quad n \to \infty . \end{aligned}$$

The argument for type II errors is similar:

$$\begin{aligned} E_Q(1 - \phi_n(\underline{X})) &= \int (1 - \phi_n(\underline{x}))\, q_n(\underline{x})\, d\mu(\underline{x}) \\ &= \int (1 - \phi_n(\underline{x}))\, q_n^{1/2}(\underline{x})\, q_n^{1/2}(\underline{x})\, d\mu(\underline{x}) \\ &\le k_n^{1/2} \int (1 - \phi_n(\underline{x}))\, p_n^{1/2}(\underline{x})\, q_n^{1/2}(\underline{x})\, d\mu(\underline{x}) \\ &= k_n^{1/2} \rho(P^n, Q^n) \\ &= k_n^{1/2} [\rho(P,Q)]^n \end{aligned}$$

$$\to 0 \quad \text{as} \quad n \to \infty . \qquad \square$$

2. Conditional Tests; Permutation Methods

Outline

2.1.1. Notation. Consider testing

$$H : \theta \in \Theta_0 \qquad \text{versus} \qquad K : \theta \in \Theta_1$$

where $X \sim P_\theta$, for some $\theta \in \Theta = \Theta_0 + \Theta_1$ is observed.
Let ϕ denote a critical function.

2.1.2. Definition. ϕ is *unbiased* if $\beta_\phi(\theta) \geq \alpha$ for all $\theta \in \Theta_1$ and $\beta_\phi(\theta) \leq \alpha$ for all $\theta \in \Theta_0$.
ϕ is *similar on the boundary* (SOB) if

$$\beta_\phi(\theta) = \alpha \qquad \text{for all} \quad \theta \in \overline{\Theta_0} \cap \overline{\Theta_1} \equiv \Theta_B \ .$$

2.1.3. Remarks. 1. If ϕ is UMP level α test, then ϕ is unbiased.
2. If ϕ is unbiased and $\beta_\phi(\theta)$ is continuous for $\theta \in \Theta$, then ϕ is SOB.

Proofs: 1. Compare with $\phi_0 \equiv \alpha$.
2. Let θ_n 's in Θ_0 converge to $\theta_0 \in \Theta_B$. Then $\beta_\phi(\theta_0) = \lim_n \beta_\phi(\theta_n) \leq \alpha$. Likewise $\beta_\phi(\theta_0) \geq \alpha$. Hence $\beta_\phi(\theta_0) = \alpha$. □

2.1.4. Definition. A *uniformly most powerful unbiased* level α test is a test ϕ_0 for which

$$E_\theta \phi_0 \geq E_\theta \phi \qquad \text{for all} \quad \theta \in \Theta_1$$

and for all unbiased level α tests ϕ .

2.1.5. Lemma. If $\mathbf{P} = \{P_\theta : \theta \in \Theta\}$ is such that $\beta_\phi(\theta)$ is continuous for all test functions ϕ , then if ϕ_0 is UMP SOB for H versus K and if ϕ_0 is level α for H versus K , then ϕ_0 is UMPU for H versus K .

Proof. The unbiased tests are a subset of the SOB tests by (2) of the remark. Since ϕ_0 is UMP SOB, it is thus at least as powerful as any unbiased test. But ϕ_0 is unbiased; since its power is $\geq$ to that of the SOB test $\phi \equiv \alpha$, and since it is level α . Thus ϕ_0 is UMPU.

□

2.1.6. Remark. For a multiparameter exponential family with densities

$$\frac{dP_\theta}{d\mu}(x) = c(\theta)\exp\left(\sum \theta_j T_j(x)\right),$$

the power function $\beta_\phi(\theta)$ is continuous in θ for all ϕ.

Proof. Apply theorem 9 of chapter 2 of Lehmann with $\phi \equiv 1$ to find that $c(\theta)$ is continuous; then apply it again with ϕ denoting an arbitrary critical function. □

2.2. Application to one - parameter exponential families.

2.2.1. The hypotheses.

Consider

$$p_\theta(\underline{x}) = c(\theta)\exp(\theta T(\underline{x}))\, h(\underline{x})$$

with respect to a $\sigma -$ finite measure μ on some subset of R^n.

Problems: Test:

(1)	$H_1 : \theta \le \theta_0$	versus	$K_1 : \theta > \theta_0$.
(2)	$H_2 : \theta \le \theta_1$ or $\theta \ge \theta_2$	versus	$K_2 : \theta_1 < \theta < \theta_2$.
(3)	$H_3 : \theta_1 \le \theta \le \theta_2$	versus	$K_3 : \theta < \theta_1$ or $\theta_2 < \theta$.
(4)	$H_4 : \theta = \theta_0$	versus	$K_4 : \theta \ne \theta_0$.

2.2.2. Theorem.

(1) The test ϕ_1 with $E_{\theta_0}\phi(T) = \alpha$ given by

$$\phi_1(T(\underline{x})) = \begin{cases} 1 & \text{if } T(\underline{x}) > c \\ \gamma & \text{if } T(\underline{x}) = c \\ 0 & \text{if } T(\underline{x}) < c \end{cases}$$

is UMP for H_1 versus K_1.

(2) The test ϕ_2 with $E_{\theta_i}\phi_2(T) = \alpha$, $i = 1, 2$ given by

$$\phi_2(T(\underline{x})) = \begin{cases} 1 & \text{if } c_1 < T(\underline{x}) < c_2 \\ \gamma_i & \text{if } T(\underline{x}) = c_i \\ 0 & \text{if else} \end{cases}$$

is UMP for H_2 versus K_2.

(3) The test ϕ_3 with $E_{\theta_i}\phi_3(T) = \alpha$, $i = 1, 2$ given by

$$\phi_3(T(\underline{x})) = \begin{cases} 1 & \text{if } T(\underline{x}) < c_1 \text{ or } T(\underline{x}) > c_2 \\ \gamma_i & \text{if } T(\underline{x}) = c_i \\ 0 & \text{if else} \end{cases}$$

is UMPU for H_3 versus K_3 .

(4) The test ϕ_4 with $E_{\theta_0}\phi_4(T) = \alpha$ and $E_{\theta_0}T\,\phi_4(T) = \alpha\, E_{\theta_0}T$ given by

$$\phi_4(T(\underline{x})) = \begin{cases} 1 & \text{if } T(\underline{x}) < c_1 \text{ or } T(\underline{x}) > c_2 \\ \gamma_i & \text{if } T(\underline{x}) = c_i \\ 0 & \text{if else} \end{cases}$$

is UMPU for H_4 versus K_4 . Furthermore, if T is symmetrically distributed about a under θ_0 , then $E_{\theta_0}\phi_4(T) = \alpha$, $c_2 = 2a - c_1$ and $\gamma_1 = \gamma_2$ determine the constants. The characteristic behavior of the power functions of these four tests is as follows:

Proof. (1) and (2) were proved earlier using the NP lemma (via MLR) and its generalized version respectively.
For (3), see page 126 in Lehmann TSH.
(4). We need only consider tests $\phi(\underline{x}) = \psi(T(\underline{x}))$ based on the sufficient statistic T , whose distribution is of the form $p_\theta(t) = c(\theta)e^{\theta t}$ with respect to some $\sigma -$ finite measure ν . Since all power functions are continuous in the case of an exponential family, we must have that any unbiased ψ satisfies $\alpha = \beta_\psi(\theta_0) = E_{\theta_0}\psi(T)$ and has a minimum at θ_0 .

But by theorem 9, chapter 2, TSH, β_ψ is differentiable, and can be differentiated under the integral sign; hence

$$\begin{aligned} \beta'_\psi(\theta) &= \frac{d}{d\theta}\int \psi(t)\,c(\theta)\exp(\theta t)\,d\nu(t) \\ &= \frac{c'(\theta)}{c(\theta)}\,E_\theta\psi(T) + E_\theta T\psi(T) \\ &= (-E_\theta T)\,E_\theta\psi(T) + E_\theta(T\psi(T)) \end{aligned}$$

since, with $\psi = \alpha$, $0 = \beta'_{\psi=\alpha} = \dfrac{c'(\theta)}{c(\theta)} + E_\theta T$.

Thus

$$0 = \beta'_\psi(\theta_0) = E_\theta(T\psi(T)) - \alpha\, E_\theta T\ .$$

Thus any unbiased test $\psi(T)$ satisfies the two conditions of the statement of our theorem. We will apply the generalized NP lemma to show that ϕ as given is UMPU.

Let

$$M \equiv \{(E_{\theta_0}\psi(T), E_{\theta_0}T\psi(T)) : \psi(T) \text{ is a critical function}\} .$$

Then M is convex and contains $\{(u, uE_{\theta_0}T) : 0 < u < 1\}$. Also M contains points (α, v) with $v > \alpha E_{\theta_0}T$; since, by problem 18 of chapter 3, Lehmann, TSH, there exist test (UMP one-sided ones) having $\beta'(\theta_0) > 0$. Likewise M contains points (α, v) with $v < \alpha E_{\theta_0}T$. Hence $(\alpha, \alpha E_{\theta_0}T)$ is an interior point of M.

Thus, by the generalized NP lemma (iv), there exist k_1, k_2 such that

$$\psi(t) = \begin{cases} 1 & \text{when } c(\theta_0)(k_1 + k_2 t)e^{\theta_0 t} < c(\theta')e^{\theta' t} \\ 0 & \text{when } c(\theta_0)(k_1 + k_2 t)e^{\theta_0 t} > c(\theta')e^{\theta' t} \end{cases}$$

(o)

$$= \begin{cases} 1 & \text{when } a_1 + a_2 t < e^{bt} \\ 0 & \text{when } a_1 + a_2 t > e^{bt} \end{cases}$$

having the property that it maximizes $E_{\theta'}\psi(T)$. But the region (o) is either one - sided or else the complement of an interval. By theorem 3.1.6 it can't be one - sided (since one - sided tests have strictly monotone power functions violating $\beta'(\theta_0) = 0$). Thus

$$\psi(T) = \begin{cases} 1 & \text{if } T < c_1 \text{ or } T > c_2 \\ 0 & \text{if } c_1 < T < c_2 \end{cases} .$$

Since this test does not depend on $\theta' \neq \theta_0$, it the UMP (within the class of level α test having $\beta'(\theta_0) = 0$) test of H_4 versus K_4. Since $\psi_\alpha \equiv \alpha$ is in this class, ψ is unbiased.

And this class of tests includes the unbiased tests. Hence ψ is UMPU.

If T is distributed symmetrically about some point a under θ_0, then any test ψ symmetric about a that satisfies $E_{\theta_0}\psi(T) = \alpha$ will also satisfy

$$E_{\theta_0}T\psi(T) = E_{\theta_0}(T-a)\psi(T) + aE_{\theta_0}\psi(T)$$
$$= 0 + a\,\alpha = \alpha E_{\theta_0}T$$

automatically. □

2.3. UMPU tests for families with nuisance parameter via conditioning

2.3.1. Definition. Let T be sufficient for $\mathbf{P}_B \equiv \{P_\theta : \theta \in \Theta_B\}$, and let $\mathbf{P}^T \equiv \{P_\theta^T : \theta \in \Theta_B\}$. A test function ϕ is said to have *Neyman structure with respect to* T if

$$E(\phi(X)\,|\,T) \;=\; \alpha \qquad \text{a.s. } \; \mathbf{P}^T \, .$$

2.3.2. Remark. If ϕ has Neyman structure with respect to T, then ϕ is SOB.

Proof. $E_\theta \phi(X) \;=\; E_\theta E(\phi(X)\,|\,T) \;=\; E_\theta \alpha \;=\; \alpha$ for all $\theta \in \Theta_B$. □

2.3.3. Theorem. Let X be a random variable with distribution $P_\theta \in \mathbf{P} \;=\; \{P_\theta : \theta \in \Theta\}$, and let T be sufficient for $\mathbf{P}_B \;=\; \{P_\theta : \theta \in \Theta_B\}$. Then all SOB tests have Neyman structure wrt T iff the family of distributions $\mathbf{P}^T \;\equiv\; \{P_\theta^T : \theta \in \Theta_B\}$ is boundedly complete: i.e. if $E_P h(T) \;=\; 0$ for all $P \in \mathbf{P}^T$ with h bounded, then $h \;=\; 0$ a.e. $\mathbf{P}^T$.

2.3.4. Remark. Suppose that:

- A All critical functions ϕ have continuous power functions β_ϕ ;
- B T is sufficient for $\mathbf{P}_B \;=\; \{P_\theta : \theta \in \Theta_B\}$ and $\mathbf{P}^T \;\equiv\; \{P_\theta^T : \theta \in \Theta_B\}$ is boundedly complete.

(Remark 2.1.6 says that A is always true for exponential families $p_\theta(x) \;=\; c(\theta)\exp(\sum \theta_j T_j(x))$; and theorem ?? allows us to check B for these same families.) Then all unbiased tests are SOB and all SOB tests have Neyman structure. Thus if we can find a UMP Neyman structure test ϕ_0 and we can show that ϕ_0 is unbiased, then ϕ_0 is UMPU. Why is it easier to find UMP Neyman structure tests? Neyman structure tests are characterized by having conditional probability of rejection equal to α on each surface $T = t$. But the distribution on each such surface is independent of $\theta \in \Theta_B$ because T is sufficient for $\mathbf{P}^T$. Thus the problem has been reduced to testing a one parameter hypothesis for each fixed value of t ; and in many problems we can easily find the most powerful test of this simple hypothesis.

Proof of theorem 2.3.3. Suppose that $\mathbf{P}^T$ is boundedly complete. Let ϕ be a SOB level α test; and define $\psi(T) \;\equiv\; E(\phi(X)\,|\,T)$. Now

$$\begin{aligned} E_\theta(\psi(T) - \alpha) \;&=\; E_\theta(E(\phi(X)\,|\,T)) \;-\; \alpha \\ &=\; E_\theta\phi(X) \;-\; \alpha \;=\; 0 \quad \text{for all } \; \theta \;\in\; \Theta_B \, , \end{aligned}$$

and since $\psi(T) \;-\; \alpha$ is bounded, the bounded completeness of $\mathbf{P}^T$ implies $\psi(T) \;=\; \alpha$ a.e. $\mathbf{P}^T$. Hence $\alpha \;=\; \psi(T) \;=\; E(\phi(X)\,|\,T)$ a.e. $\mathbf{P}^T$, and ϕ has Neyman structure with respect to T .

Now suppose that all SOB tests have Neyman structure. Assume $\mathbf{P}^T$ is *not* boundedly complete. Then there exists h such that $|h| \leq$ some M with $E_\theta h(T) \;=\; 0$ for all $\theta \in \Theta_B$ and $h(T) \;\neq\; 0$ with probability > 0 for some $\theta_0 \in \Theta_B$. Define $\phi(T) \;\equiv\; c\,h(T) \;+\; \alpha$ where $c \;\equiv\; \{\alpha \bigwedge (1-\alpha)\}/M$. Then $0 \;\leq\; \phi(T) \;\leq\; 1$ so ϕ is a critical function, and $E_\theta\phi(T) \;=\; \alpha$ for all $\theta \in \Theta_B$, so that ϕ is SOB. But $E(\phi(T)\,|\,T) = \phi(T) \neq \alpha$ with probability > 0 for the above $\theta_0 \in \Theta_B$, so ϕ does not have Neyman structure. This is a contradiction, and hence it follows that indeed $\mathbf{P}^T$ is boundedly complete. □

2.3.5. Examples.

Example 1. (One - sample; $N(\mu, \sigma^2)$ with σ^2 as nuisance parameter).

Example 2. (Comparing two Poisson distributions).

Example 3. (Comparing two Binomial distributions).

Example 4. (Comparing two Normal means when variances are equal).

Example 5. (Paired normals with nuisance shifts).

2.4. Application to general exponential families; k − parameter.

2.4.1. The hypotheses: Consider the exponential family $\mathbf{P} = \{P_{\theta,\xi}\}$ given by

$$p_{\theta,\xi}(\underline{x}) = c(\theta, \xi)\exp[\theta U(\underline{x}) + \sum_{i=1}^{k} \xi_i T_i(\underline{x})]$$

with respect to a $\sigma-$ finite μ on some subset of R^n where Θ is convex, has dimension $k+1$, and contains interior points θ_i, $i = 0, 1, 2$.

Problems: Test

(1)	$H_1 : \theta \leq \theta_0$	versus	$K_1 : \theta > \theta_0$.
(2)	$H_2 : \theta \leq \theta_1$ or $\theta \geq \theta_2$	versus	$K_2 : \theta_1 < \theta < \theta_2$.
(3)	$H_3 : \theta_1 \leq \theta \leq \theta_2$	versus	$K_3 : \theta < \theta_1$ or $\theta_2 < \theta$.
(4)	$H_4 : \theta = \theta_0$	versus	$K_4 : \theta \neq \theta_0$.

2.4.2. Theorem. The following are UMPU tests for the hypothesis testing problems 1 - 4 respectively:

(1) $$\phi_1(\underline{x}) = \begin{cases} 1 & \text{if } U > c(t) \\ \gamma(t) & \text{if } U = c(t) \\ 0 & \text{if } U < c(t) \end{cases}$$

where $E_{\theta_0}(\,\phi_1(U)\,|\,T = t) = \alpha$.

(2) $$\phi_2(\underline{x}) = \begin{cases} 1 & \text{if } c_1(t) < U < c_2(t) \\ \gamma_i & \text{if } U = c_i(t) \\ 0 & \text{if else} \end{cases}$$

where $E_{\theta_i}(\phi_2(U)\,|\,T = t) = \alpha$, $i = 1, 2$.

(3) $$\phi_3(\underline{x}) = \begin{cases} 1 & \text{if } U < c_1(t) \text{ or } U > c_2(t) \\ \gamma_i & \text{if } U = c_i(t) \\ 0 & \text{if else} \end{cases}$$

where $E_{\theta_i}(\phi_3(U) \,|\, T = t) = \alpha \,, \; i = 1, 2$.

(4) $$\phi_4(\underline{x}) = \begin{cases} 1 & \text{if } U < c_1(t) \text{ or } U > c_2(t) \\ \gamma_i & \text{if } U = c_i(t) \\ 0 & \text{if else} \end{cases}$$

where $E_{\theta_0}(\phi_4(U) \,|\, T = t) = \alpha$ and $E_{\theta_0}(U \phi_4(U) \,|\, T = t) = \alpha \, E_{\theta_0}(U \,|\, T = t)$.

2.4.3. Remarks. See TSH, theorem 5.1, page 161.

A. If $V = h(U, T)$ is increasing in U for each fixed t and is independent of T on Θ_B , then

$$\phi(\underline{x}) = \begin{cases} 1 & \text{if } V > c \\ \gamma & \text{if } V = c \\ 0 & \text{if } V < c \end{cases}$$

is UMPU in (1).

B. If $h \equiv h(U, T) = a(t)U + b(t)$ with $a(t) > 0$, then the second constraint in (4) becomes

$$E_{\theta_0}\left(\frac{V - b(t)}{a(t)} \, \phi \,|\, T = t \right) = \alpha E_{\theta_0}\left(\frac{V - b(t)}{a(t)} \,|\, T = t \right)$$

or $E_{\theta_0}(V \phi \,|\, T = t) = \alpha E_{\theta_0}(V \,|\, T = t)$, and if this V is independent of T on the boundary, then the test is unconditional.

2.5. Permutation tests.

2.5.1. The two - sample problem:

Consider testing

$$H_c : \; X_1, \cdots, X_m, Y_1, \cdots, Y_n \quad \text{are iid with df} \quad F \in \mathbf{F}_c$$

where $\mathbf{F}_c \equiv$ {all continuous df's F} versus

$$K_1 : \; X_1, \cdots, X_m, Y_1, \cdots, Y_n \quad \text{have joint density function} \quad h \, .$$

We seek a most powerful similar test: ϕ is *similar* if

(1) $$E_{(F,F)} \phi(\underline{X}, \underline{Y}) = \alpha \qquad \text{for all} \quad F \in \mathbf{F}_c \, .$$

But if $\underline{Z} \equiv (Z_1, \cdots, Z_N)$ with $N \equiv m + n$ denotes the ordered values of the combined sample $X_1, \cdots, X_m, Y_1, \cdots, Y_n$, then when H_c is true, $\underline{Z}$ is sufficient and complete. Hence (1) holds iff (by theorem 6.2.3.3)

$$E(\phi(\underline{X}, \underline{Y}) \,|\, \underline{Z} = \underline{z}) = \alpha \qquad \text{for a.e.} \quad \underline{z} = (z_1, \cdots, z_N)$$

(2) $$= \sum_{\pi \in \Pi} \phi(\pi \underline{z}) \frac{1}{N!} = \sum_{\underline{z}'} \phi(\underline{z}') \frac{1}{N!}$$

where the sum is over all $N!$ permutations $\underline{z}'$ of $\underline{z}$. Thus if $\alpha = I/N!$, then any test which is performed conditionally on $\underline{Z} = \underline{z}$ and rejects for exactly I of the $N!$ permutations $\underline{z}'$ of $\underline{z}$ is a level α similar test; moreover (2) says that any level α similar test is of this form.

2.5.2. Definition. Tests satisfying (2) are called *permutation tests.* (Thus a test of H_c versus K_1 is similar iff it is a permutation test.)

We now need to find a most powerful permutation test by maximizing the conditional power. But

$$E_h(\phi(\underline{X}\,,\underline{Y}) \mid \underline{Z} = \underline{z}) \;=\; \sum_{\underline{z}'} \phi(\underline{z}') \frac{h(\underline{z}')}{\sum h(\underline{z}'')}\,,$$

Since the conditional densities under the composite null hypothesis and under the simple alternative h are

$$p_0(\underline{z}' \mid \underline{z}) \;=\; \frac{1}{N!} \qquad \text{and} \qquad p_1(\underline{z}' \mid \underline{z}) \;=\; \frac{h(\underline{z}')}{\sum h(\underline{z}'')}\,, \qquad \underline{z}' \in \{\pi\underline{z} : \pi \in \Pi\,,$$

the conditional power is maximized by rejecting for large values of

$$\frac{p_1(\underline{z}' \mid \underline{z})}{p_0(\underline{z}' \mid \underline{z})} \;=\; K_{\underline{z}}\, h(\underline{z}') \qquad \text{with} \quad K_{\underline{z}} \;=\; \frac{N!}{\sum h(\underline{z}'')}\,.$$

Thus, at level $\alpha = I/N!$ we reject if

(3) $$h(\underline{z}') \;>\; c_{\underline{z}}$$

where $c_{\underline{z}}$ is chosen so that we reject for exactly I of the $N!$ permutations $\underline{z}'$ of $\underline{z}$; or else we use a randomized version of such a test.

2.5.3. Example. Suppose now that we specify a particular alternative:

$$K_1 : \begin{array}{l} X_1\,, \cdots, X_m \quad \text{are iid } N(\theta_1\,, \sigma^2) \\ Y_1\,, \cdots, Y_n \quad \text{are iid } N(\theta_2\,, \sigma^2) \end{array}$$

where $\theta_1 < \theta_2$ and σ^2 are fixed constants. Then the similar test of H_c that is most powerful against this simple K_1 rejects H for those permutations $\underline{z}'$ of $\underline{z}$ which lead to large values of

$$(2\pi\sigma^2)^{-N/2} \exp\{-\frac{1}{2\sigma^2}[\sum_1^m (X_i - \theta_1)^2 \;+\; \sum_1^n (Y_j - \theta_2)^2]\}$$

or small values of

$$\sum_1^m (X_i - \theta_1)^2 \;+\; \sum_1^n (Y_j - \theta_2)^2$$

$$= \sum_1^m X_i^2 \;+\; \sum_{j=1}^n Y_j^2 \;+\; m\theta_1^2 \;+\; n\theta_2^2 \;-\; 2\theta_1 \sum_{i=1}^m X_i \;-\; 2\theta_2 \sum_{j=1}^n Y_j\,,$$

or large values of

$$\theta_1 \sum_{i=1}^{m} X_i \; + \; \theta_2 \sum_{j=1}^{n} Y_j \; ,$$

or large values of

$$\theta_1 \sum_{i=1}^{m} X_i \; + \; \theta_2 \sum_{j=1}^{n} Y_j \; - \; \frac{m\theta_1 + n\theta_2}{N} \left(\sum X_i \; + \; \sum Y_j \right)$$

$$= \; \frac{mn}{N} (\theta_2 - \theta_1)(\bar{Y} \; - \; \bar{X}) \; ,$$

or large values of

$$\bar{Y} \; - \; \bar{X}$$

or large values of

$$\theta_1 \sum_{i=1}^{m} X_i \; + \; \theta_2 \sum_{j=1}^{n} Y_j \; - \; \theta_1 \left(\sum X_i \; + \; \sum Y_j \right) \; = \; (\theta_2 \; - \; \theta_1) \sum_{j=1}^{n} Y_j \; ,$$

or large values of

$$\frac{\sqrt{\frac{mn}{N}} (\bar{Y} \; - \; \bar{X})}{\sqrt{\frac{1}{N-2} \left\{ \sum Z_i^2 \; - \; \frac{(\sum Z_i)^2}{N} \; - \; \frac{mn}{N} (\bar{Y} - \bar{X})^2 \right\}}}$$

$$= \; \frac{\sqrt{\frac{mn}{N}} (\bar{Y} \; - \; \bar{X})}{\sqrt{\frac{1}{N-2} \left\{ \sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2 \right\}}} \; \equiv \; \tau \; .$$

Thus the most powerful similar test of H_c versus K_1 is

$$\phi(\underline{z}') \; = \; \begin{cases} 1 & \text{if} \quad \tau \; > \; c_\alpha(\underline{z}) \\ 0 & \text{if} \quad \tau \; < \; c_\alpha(\underline{z}) \end{cases}$$

where $c_\alpha(\underline{z})$ is chosen so that exactly $\alpha N!$ of the permutations $\underline{z}'$ lead to rejection (if this is possible; if not we can use a randomized test). But we know that τ on at most $\binom{N}{m}$ distinct values according to each of the $\binom{N}{m}$ assignments $\underline{z}_c$ of m of the z_i 's to be X_i 's. Thus

$$\phi(\underline{z}_c) \; = \; \begin{cases} 1 & \text{if} \quad \tau(\underline{z}_c) \; > \; c_\alpha(\underline{z}) \\ 0 & \text{if} \quad \tau(\underline{z}_c) \; < \; c_\alpha(\underline{z}) \end{cases} \tag{4}$$

where $c_\alpha(\underline{z})$ is chosen so that exactly $\alpha \binom{N}{m}$ of the assignments $\underline{z}_c$ of m of the z_i 's to be X_i 's leads to rejection.

Since the test (4) does not depend on which $\theta_1 < \theta_2$ or σ^2 we started with, the test is actually a UMP similar test of H_c versus $K \equiv \cup_{\theta_1 < \theta_2, \sigma^2} K_1$; i.e. different normal distributions with $\theta_1 < \theta_2, \sigma^2$ unknown.

Example. Suppose that $(X_1, X_2) = (56, 72)$, $(Y_1, Y_2, Y_3) = (47, 68, 86)$. Thus $\bar{X} = 64$, $\bar{Y} = 67$, $\bar{Y} - \bar{X} = 3$. Here

$$\underline{Z} = (47, 56, 68, 72, 86)$$

and $\binom{5}{2} = \frac{5!}{2!3!} = 10$. (Note that $5! = 120$.)

combination	47	56	68	72	86	$\bar{Y} - \bar{X}$
1	Y	Y	Y	X	X	– 22. 0
2	Y	Y	X	Y	X	– 18. 7
3	Y	X	Y	Y	X	– 8. 7
4	X	Y	Y	Y	X	– 1. 2
5	Y	Y	X	X	Y	– 7. 0
6	Y	X	Y	X	Y	3. 0
7	X	Y	Y	X	Y	10. 5
8	Y	X	X	Y	Y	6. 3
9	X	Y	X	Y	Y	13. 8
10	X	X	Y	Y	Y	23. 8

Note that $\binom{20}{10} = 184,756$, and, by Stirling's formula -- $m! \sim \sqrt{2\pi m}(m/e)^m$, that

$$\binom{2m}{m} \sim \frac{1}{\sqrt{\pi m}} 2^{2m} \qquad \text{as } m \to \infty ,$$

so the exact permutation test is difficult computationally for all but small sample sizes. But *sampling* from the permutation distribution is always possible.

2.5.4. Remarks. We will call the present test "reject if $\tau > c_\alpha(\underline{z})$ " the *permutation t - test*; it is the UMP similar test of H_c versus K specified above. If we consider the smaller null hypothesis

$$H_G : X_1, \cdots, X_m, Y_1, \cdots, Y_n \text{ i.i.d. } N(\theta, \sigma^2) \text{ with } \theta, \sigma^2 \text{ unknown,}$$

then we recall that the *classical* $t -$ *test "reject if* $\tau > t_{m+n-2,\alpha}$ *" is the UMPU test of* H_G *versus* K .

The classical $t -$ test has greater power than the permutation $t -$ test for H_G ; but it is not a similar test of H_c .

If we could show that for a.e. $\underline{z}$ the numbers

$$c_\alpha(\underline{z}) \qquad \text{and} \qquad t_{m+n-2,\alpha}$$

where just about equal, then the classical $t -$ *test* and the permutation $t -$ test would be almost identical.

2.4.5. Theorem. If $F \in \mathbf{F}_c$ has $E_F|X|^2 < \infty$ and if $0 < \liminf \frac{m}{N} \leq \limsup \frac{m}{N} < 1$, then

$$c_\alpha(\underline{z}) \to z_\alpha$$

where $P(N(0,1) > z_\alpha) = \alpha$. Since we also know that $t_{m+n-2,\alpha} \to z_\alpha$, it follows that $c_\alpha(\underline{z}) - t_{m+n-2,\alpha} \to 0$.

Proof. Let an urn contain balls numbered $z_1, \cdots, z_N$. Let $Y_1, \cdots, Y_n$ denote the numbers on n balls drawn without replacement. Let $\bar{z} = N^{-1}\sum_{i=1}^N z_i$, $\sigma_z^2 = N^{-1}\sum_{i=1}^N (z_i - \bar{z})^2$, $m \equiv N - n$. Then

$$E\bar{Y} = \bar{z} \qquad \text{and} \qquad \sigma_N^2 \equiv Var(\bar{Y}) = (1 - \frac{n-1}{N-1})\frac{\sigma_z^2}{n}.$$

Moreover, by the Wald - Wolfowitz - Noether - Hájek finite sampling CLT

$$\frac{(\bar{Y} - \bar{z})}{\sigma_N} \to_d N(0,1)$$

as long as the Noether condition

$$\eta_N \equiv \frac{\max_{1 \leq i \leq N} |z_i - \bar{z}|^2}{\sum_{i=1}^N |z_i - \bar{z}|^2} \to 0$$

holds.

Now rewrite the permutation t – statistic τ : note that

$$\begin{aligned} \bar{Y} - \bar{X} &= \bar{Y} - \frac{1}{m}\sum X_i - \frac{1}{m}\sum Y_i + \frac{n}{m}\bar{Y} \\ &= \frac{N}{m}(\bar{Y} - \bar{z}), \end{aligned}$$

and hence

$$\sqrt{\frac{mn}{N}}(\bar{Y} - \bar{X}) = \sqrt{\frac{mn}{N}}\frac{N}{m}(\bar{Y} - \bar{z}) = \sqrt{\frac{N}{N-1}}\sqrt{\frac{N-1}{m}}\sqrt{n}(\bar{Y} - \bar{z}).$$

Thus

$$\begin{aligned} \tau &= \frac{\sqrt{\frac{mn}{N}}(\bar{Y} - \bar{X})}{\sqrt{\frac{1}{N-2}\{\sum Z_i^2 - \frac{(\sum Z_i)^2}{N} - \frac{mn}{N}(\bar{Y} - \bar{X})^2\}}} \\ &= \frac{\sqrt{\frac{N}{N-1}}\frac{\bar{Y} - \bar{z}}{\sqrt{\frac{1}{n}(1 - \frac{n-1}{N-1})}}}{\sqrt{\frac{N}{N-2}\sigma_z^2 - \frac{1}{N-2}\frac{N}{N-1}\frac{(\bar{Y} - \bar{z})^2}{\frac{1}{n}(1 - \frac{n-1}{N-1})}}} \end{aligned}$$

$$= \sqrt{\frac{N-2}{N-1}} \frac{(\bar{Y} - \bar{z})/\sigma_N}{\sqrt{1 - \frac{1}{N-1} \frac{(\bar{Y} - \bar{z})^2}{\sigma_N^2}}}$$

$$\to_d \; 1 \cdot \frac{Z}{\sqrt{1 - 0\, Z^2}} = Z \sim N(0,1)$$

if

$$\frac{\bar{Y} - \bar{z}}{\sigma_N} \to_d Z \sim N(0,1) \,;$$

in probability or, better yet, almost surely; i.e. if

$$P(\frac{\bar{Y} - \bar{z}}{\sigma_N} \le t \,|\, \underline{Z} = \underline{z}) \to \Phi(t)$$

in probability or almost surely. But this holds under the present hypotheses in view of the following finite - sampling CLT.

2.4.6. Theorem. (Wald - Wolfowitz - Noether - Hájek finite - sampling limit theorem). If $0 < \liminf \frac{m}{N} \le \limsup \frac{m}{N} < 1$, then

$$\frac{\bar{Y} - \bar{z}}{\sigma_N} \to_d Z \sim N(0,1) \quad \text{as} \quad N \to \infty$$

if and only if

$$\eta_N \equiv \frac{\max_{1 \le i \le n} |z_i - \bar{z}|^2}{\sum_{i=1}^{N} (z_i - \bar{z})^2} \to 0 \quad \text{as} \quad N \to \infty \,.$$

Moreover

$$\sup_t |P(\frac{\bar{Y} - \bar{z}}{\sigma_N} \le t) - \Phi(t)| \le 5(\frac{N}{m \wedge n} \eta_N)^{1/4} \quad \text{for all} \quad N \ge 1 \,.$$

Proof. See Hájek, *Ann. Math. Statist.* **32,** 506 - 523. For still better rates under stronger conditions, see Bolthausen (1984). □

To complete the proof of theorem 2.5.5, it suffices to show that

(5) $$\eta_N \to 0 \quad a.s.$$

But, even under the alternative hypothesis $F \ne G$ and $E_F|X|^2 < \infty$, $E_G|Y|^2 < \infty$,

$$\frac{1}{N} \sum_{i=1}^{N} (Z_i - \bar{Z})^2 = \frac{1}{N} \{ \sum_{i=1}^{N} Z_i^2 - N \bar{Z}^2 \}$$

$$= \frac{(m-1)}{N} S_X^2 + \frac{(n-1)}{N} S_Y^2$$

$$\rightarrow_{a.s.} \ \lambda\sigma_X^2 \ + \ (1-\lambda)\sigma_Y^2 \ \geq \ \min\{\sigma_X^2\,, \sigma_Y^2\} \ > \ 0$$

for any subsequence $N \rightarrow \infty$ for which $\lambda_N \equiv m/N \rightarrow \lambda$, and hence the denominator of η_N (divided by N) has a positive limit inferior almost surely. To see that the numerator converges almost surely to zero, first recall that $\max_{1 \leq i \leq n}|X_i|/n \rightarrow_{a.s.} 0$ iff $E_F|X_1| < \infty$. Hence $\max_{1 \leq i \leq n}|X_i|^2/n \rightarrow_{a.s.} 0$ iff $E_F|X_1|^2 < \infty$. Thus we rewrite the numerator divided by N as

$$\begin{aligned} \frac{1}{N}\max_{i \leq N}|Z_i - \bar{Z}|^2 &\leq \frac{2}{N}\{\max_{i \leq N}|Z_i|^2 + \bar{Z}^2\} \\ &\leq \frac{2}{N}\{\max[\max_{i \leq m}|X_i|^2, \max_{j \leq n}|Y_j|^2] + (\frac{m}{N}\bar{X} + \frac{n}{N}\bar{Y})^2\} \\ &\leq 2\max\{\frac{1}{m}\max_{i \leq m}|X_i|^2, \frac{1}{n}\max_{j \leq n}|Y_j|^2\} + \frac{2}{N}[\frac{m}{N}\bar{X} + \frac{n}{N}\bar{Y}]^2\} \\ &\rightarrow_{a.s.} 0 + 0\,. \end{aligned}$$

Hence (5) holds (even under the alternative if $E_F X^2 < \infty$ and $E_G Y^2 < \infty$).

3. INVARIANT TESTS: RANK METHODS

Outline

3.1. Notation and basic results.

3.1.1. Notation and Assumptions.

Let $(\mathbf{X}, \mathbf{A}, P_\theta)$ be a probability space for all $\theta \in \Theta$, and suppose $\theta \neq \theta'$ implies $P_\theta \neq P_{\theta'}$.
We observe $X \sim P_\theta$.
Suppose that $g : \mathbf{X} \to \mathbf{X}$ is one-to-one, onto $\mathbf{X}$, and measurable, and suppose that the distribution of gX when $X \sim P_\theta$ is some $P_{\theta'} = P_{\overline{g}\theta}$; that is

(1) $$P_\theta(gX \in A) = P_{\overline{g}\theta}(X \in A) \quad \text{for all } A \in \mathbf{A},$$

or equivalently,

$$P_\theta(g^{-1}A) = P_{\overline{g}\theta}(A) \quad \text{for all } A \in \mathbf{A};$$

or equivalently

$$P_\theta(A) = P_{\overline{g}\theta}(gA) \quad \text{for all } A \in \mathbf{A}.$$

Hence

(2) $$E_\theta h(g(X)) = E_{\overline{g}\theta} h(X).$$

Suppose that $\overline{g}\Theta = \Theta$.
Let G denote a *group* of such transformations g.

We want to test $H : \theta \in \Theta_H$ versus $K : \theta \in \Theta_K$.

3.1.2. Proposition. $\overline{G}$ is a group of one-to-one transformations of Θ onto Θ and is homomorphic to G.

Proof. Suppose that $\overline{g}\theta_1 = \overline{g}\theta_2$. Then $P_{\theta_1} = P_{\theta_2}$ by (1). Thus $\theta_1 = \theta_2$ by assumption. Thus $\overline{g} \in \overline{G}$ is one-to-one.

Closure, associativity, and identity are easy.

If $X \sim P_\theta$, then $g_1 X \sim P_{\overline{g}_1\theta}$, and $(g_2 \circ g_1)X = g_2 \circ (g_1 X) \sim P_{\overline{g}_2 \circ \overline{g}_1 \theta}$, while

$(g_2 \circ g_1)X \sim P_{\overline{g_2 \circ g_1}\theta}$, so $\overline{g_2 \circ g_1} = \overline{g}_2 \circ \overline{g}_1$.
If $X \sim P_\theta$, then $g^{-1}X \sim P_{\overline{g^{-1}}\theta}$, so $g \circ g^{-1} X \sim P_{\overline{g} \circ \overline{g^{-1}}\theta}$, while $g \circ g^{-1} X = X \sim P_\theta$, so $\overline{g} \circ \overline{g^{-1}} = \overline{e}$; thus $\overline{g}^{-1} = \overline{g^{-1}}$, and $\overline{G}$ is a group. □

3.1.3. Definition. A group G of one-to-one transformations of $\mathbf{X}$ onto $\mathbf{X}$ is said to leave the decision problem H versus K *invariant* provided $\overline{g}\Theta = \Theta$ and $\overline{g}\Theta_H = \Theta_H$ for all $g \in G$.

3.2. Orbits and maximal invariants.

3.2.1. Definition. $x_1 \sim x_2 \; mod(G)$ if $x_2 = g(x_1)$ for some $g \in G$.

3.2.2. Proposition. $\sim$ is an equivalence relation.

Proof. Reflexive: $x_1 \sim x_1$ since $x_1 = e(x_1)$.
Symmetric: $g(x_1) = x_2$ implies $g^{-1}(x_2) = x_1$.
Transitive: $x_1 \sim x_2$ and $x_2 \sim x_3$ implies $x_1 \sim x_3$, since $g_1(x_1) = x_2$ and $g_2(x_2) = x_3$ implies $(g_2 \circ g_1)(x_1) = x_3$. □

3.2.3. Definition. The equivalence classes of $\sim$ are called the *orbits* of G . Thus $orbit(x) = \{g(x) : g \in G\}$.

3.2.4. Proposition. ϕ is *invariant* if and only if ϕ is constant on each orbit of G .

Proof. This follows immediately from the definitions. □

3.2.5. Definition. A measurable function $T : \mathbf{X} \to$ some R^k is a *maximal invariant for* G , or *GMI*, if T is invariant and $T(x_1) = T(x_2)$ implies $x_1 \sim x_2$. That is, T is constant on the orbits of G and takes on distinct values on distinct orbits.

3.2.6. Theorem. Let T be a GMI. Then ϕ is invariant if and only if there exists a function h such that $\phi(x) = h(T(x))$ for all $x \in \mathbf{X}$.

Proof. Suppose that $\phi(x) = h(T(x))$. Then

$$\phi(gx) = h(T(gx)) = h(T(x)) = \phi(x) ,$$

so ϕ is invariant.

On the other hand, suppose that ϕ is $G-$invariant. Then $T(x_1) = T(x_2)$ implies $x_1 \sim x_2$ implies $g(x_1) = x_2$ for some $g \in G$. Thus $\phi(x_2) = \phi(gx_1) = \phi(x_1)$; that is, ϕ is constant on the orbit. It follows that ϕ is a function of T . □

3.3. Examples

3.3.1. Translation group. Suppose $\mathbf{X} = R^n$ and $G = \{g : g\underline{x} = \underline{x} + c\underline{1} \text{ where } c \in R\}$. Then $T(\underline{x}) = (x_1 - x_n, \cdots, x_{n-1} - x_n)$ is a GMI.

Proof. Clearly T is invariant. Suppose that $T(\underline{x}) = T(\underline{x}^*)$. Then $x_i = x_i^* - (x_n^* - x_n)$ for $i = 1, \cdots, n-1$, and this holds trivially for $i = n$. Thus $\underline{x}^* = g(\underline{x}) = \underline{x} + c\underline{1}$ with $c = (x_n^* - x_n)$. □

Note that $T(\underline{x}) = (x_1 - \overline{x}, \cdots, x_n - \overline{x})$ is also a GMI, since T is invariant and $T(\underline{x}) = T(\underline{x}^*)$ implies that $x_i^* = x_i - (\overline{x} - \overline{x}^*)$, $i = 1, \cdots, n$, so $\underline{x}^* = g(\underline{x}) = x + c\underline{1}$ with $c = (\overline{x}^* - \overline{x})$.

3.3.2. Scale group. Suppose that $\mathbf{X} = \{x \in R^n : x_n \neq 0\}$, and $G = \{g : g\underline{x} = c\underline{x} \text{ where } c \in R, c \neq 0\}$. Then $T(\underline{x}) = (x_1/x_n, \cdots, x_{n-1}/x_n)$ is a GMI.

Proof. Clearly T is invariant. Suppose that $T(\underline{x}) = T(\underline{x}^*)$. Then $x_i^* = (x_n^*/x_n)x_i$ for $i = 1, \cdots, n-1$, and this holds trivially for $i = n$. Thus $\underline{x}^* = g(\underline{x}) = c\,\underline{x}$ with $c = (x_n^*/x_n)$. □

Note that $T(\underline{x}) = (x_1/\overline{x}, \cdots, x_n/\overline{x})$ is also a GMI, since T is invariant and $T(\underline{x}) = T(\underline{x}^*)$ implies that $x_i^* = (\overline{x}^*/\overline{x})x_i$, $i = 1, \cdots, n$, or $\underline{x}^* = g(\underline{x}) = c\,x$ with $c = (\overline{x}^*/\overline{x})$.

3.3.3. Orthogonal group. Suppose $\mathbf{X} = R^n$ and $G = \{g : g\underline{x} = \Gamma\underline{x},\ \Gamma \text{ an } n \times n \text{ orthogonal matrix}\}$. Then $T(\underline{x}) = \underline{x}^T\underline{x} = \sum_{i=1}^n x_i^2$ is a GMI.

Proof. $T(g\underline{x}) = \underline{x}^T\Gamma^T\Gamma\underline{x} = \underline{x}^T\underline{x} = T(\underline{x})$, so T is invariant. Suppose that $T(\underline{x}) = T(\underline{x}^*)$. Then there exists $\Gamma = \Gamma_{x,x^*}$ such that $\underline{x}^* = \Gamma\underline{x} = g(\underline{x})$. □

3.3.4. Permutation group. Suppose that $\mathbf{X} = R^n - \text{ties}$, and $G = \{g : g(\underline{x}) = \pi\underline{x} = (x_{\pi(1)}, \cdots, x_{\pi(n)})$ for some permutation $\pi = (\pi(1), \cdots, \pi(n))$ of $\{1, \cdots, n\}$. Note that $\#(G) = n!$. Then $T(\underline{x}) = (x_{(1)}, \cdots, x_{(n)}) \equiv \underline{x}_{(\cdot)}$, the vector of ordered x's is a GMI

Proof. $T(g\underline{x}) = T(\pi\underline{x}) = \underline{x}_{(\cdot)} = T(\underline{x})$, so T is invariant. Moreover, if $T(\underline{x}^*) = T(\underline{x})$, then $\underline{x}^* = \pi\underline{x}$ for some $\pi \in \Pi$, so T is maximal. □

3.3.5. Rank transformation group. Suppose that $\mathbf{X} = \{\underline{x} \in R^n : x_i \neq x_j \text{ for all } i \neq j\} = R^n - \text{ties}$, and let $G = \{g : g(x) = (f(x_1), \cdots, f(x_n)),\ f \text{ continuous and strictly increasing}\}$. Then $T(\underline{x}) \equiv \underline{r} = (r_1, \cdots, r_n)$ where r_i denotes the rank of x_i.

Proof. T is clearly invariant. If $T(\underline{x}^*) = T(\underline{x})$, then relabeling if necessary, we have a picture like:

□

3.3.6. Sign group. Suppose that $\mathbf{X} = R^n$ and that $G = \{g, e\}^n$ where $g(x) = -x$ and $e(x) = x$. Then $T(\underline{x}) = (|x_1|, \cdots, |x_n|)$ is a GMI.

3.3.7. Affine group. Suppose that $\mathbf{X} = \{x \in R^n : x_{n-1} \neq x_n\}$ and that $G = \{g : g(\underline{x}) = a\underline{x} + b\underline{1}$ with $a \neq 0$, b real$\}$. Then

$$T(\underline{x}) = (\frac{x_1 - x_n}{x_{n-1} - x_n}, \cdots, \frac{x_{n-2} - x_n}{x_{n-1} - x_n})$$

is a GMI. Note that

$$T(\underline{x}) = (\frac{x_1 - \overline{x}}{s}, \cdots, \frac{x_n - \overline{x}}{s})$$

is also a GMI (on $\mathbf{X} \equiv \{x \in R^n : s > 0\}$ where $s^2 = n^{-1}\sum_{i=1}^n (x_i - \overline{x})^2$).

Remark. Consider $G = G_2 \bigoplus G_1 =$ scale $\bigoplus$ translation $= \{g_2 \circ g_1 : g_1 \in G_1,\ g_2 \in G_2\}$. Then $Y = (x_1 - x_n, x_2 - x_n, \cdots, x_{n-1} - x_n)$ is a G_1MI. In the space of the G_1MI we have $Z = (y_1/y_{n-1}, \cdots, y_{n-2}/y_{n-1})$ is a G_2MI. Thus Z is the GMI.
If this works, it's OK; see theorem 2 on TSH page 218.
But it doesn't always work. When $G = G_1 \bigoplus G_2$, it does work if G_1 is a normal subgroup of G. [Recall that G_1 is a normal subgroup of G if and only if $gG_1g^{-1} = G_1$ for all $g \in G$.]

3.3.8. The signed rank transformation group. Suppose that $\mathbf{X} = \{\underline{x} \in R^n : x_i \neq x_j$ for all $i \neq j\} = R^N -$ ties as in 3.3.5 (but with N instead of n), but now let $G = \{g : g(x) = (f(x_1), \cdots, f(x_N))$, f is odd, continuous, and strictly increasing$\}$. Then $T(\underline{x}) \equiv (\underline{r}, \underline{s}) = (r_1, \cdots, r_m, s_1, \cdots, s_n)$ where $r_1, \cdots, r_m$ denote the rank of $|x_{i_1}|, \cdots, |x_{i_m}|$ among $|x_1|, \cdots, |x_N|$, and $s_1, \cdots, s_n$ denote the rank of $|x_{j_1}|, \cdots, |x_{j_n}|$ among $|x_1|, \cdots, |x_N|$ where $x_{i_1}, \cdots, x_{i_m} < 0 < x_{j_1}, \cdots, x_{j_n}$.
Proof. T is clearly invariant. To show maximal invariance, the picture is much as in example 3.3.5, but with the function f being odd; see Lehmann TSH pages 316-317.

3.3.9. Example. Suppose that $\mathbf{X} = \{(x_1, x_2): x_2 > 0\}$ and that $G = \{g : g(x) = (x_1 + b, x_2),\ b \in R\}$. Then $T(x) = x_2$ is a GMI.

3.3.10. Example. Suppose that $\mathbf{X} = \{(x_1, x_2, x_3, x_4) :\ x_3, x_4 > 0\}$ and that $G = \{g : g(\underline{x}) = (cx_1 + a, cx_2 + b, cx_3, cx_4), a, b \in R, c > 0\}$. Then $T(x) = x_3/x_4$ is a GMI.

3.3.11. Example. Suppose that $\mathbf{X} = \{(x_1, x_2): x_2 > 0\}$ as in example 9, but now suppose that the group $G = \{g : g(x) = (cx_1, |c|x_2) : c > 0\}$ or $G = \{g : g(x) = (cx_1, |c|x_2) : c \neq 0\}$. Then $T(x) = x_1/x_2$ is a GMI in the first case ($c > 0$), and $T(x) = |x_1|/x_2$ is a GMI in the second case ($c \neq 0$).

3.4. UMP G - invariant tests.

3.4.1. Theorem. If $T(X)$ is any $G -$ invariant function and if $\nu(\theta)$ is a $\bar{G}$MI, then the distribution of $T(X)$ depends on θ only through $\nu(\theta)$.

Proof. Suppose that $\nu(\theta_1) = \nu(\theta_2)$. Then there exists $\bar{g} \in \bar{G}$ such that $\bar{g}\theta_1 = \theta_2$. Let g be the elements of G corresponding to $\bar{g} \in \bar{G}$. Then by (1)

$$P_{\theta_1}(T(X) \in A) = P_{\theta_1}(T(gX) \in A) = P_{\bar{g}\theta_1}(T(X) \in A) = P_{\theta_2}(T(X) \in A)$$

for all $A \in \mathbf{A}$. Thus the distribution of T is a function of $\nu(\theta)$. □

3.4.2. Theorem. Suppose that H versus K is invariant under G . Let $T(X)$ and $\delta \equiv \nu(\theta)$ denote the GMI and the $\bar{G}$MI; and suppose both are real - valued. Suppose that the densities $p_\delta(t) = (dP_\delta^T/d\mu)(t)$ with respect to some $\sigma -$ finite μ have MLR in T ; and suppose that H versus K equivalent to $H_1 : \delta \le \delta_0$ versus $K_1 : \delta > \delta_0$. Then there exists a UMP G - invariant level α test of H versus K given by

$$\psi(T) = \begin{cases} 1 & \text{if } T > c \\ \gamma & \text{if } T = c \\ \gamma & \text{if } T < c \end{cases}$$

with $E_{\delta_0}\psi(T) = \alpha$.

Proof. By theorem 3.2.6, any G - invariant test ϕ is of the form $\phi = \psi(T)$. By theorem 3.4.1, the distribution of T depends only on δ . Thus our theorem for UMP tests when there is MLR completes the proof. □

3.4.3. Application 1. Tests of σ^2 for $N(\mu, \sigma^2)$. Let $X_1, \cdots, X_n$ be iid $N(\mu, \sigma^2)$. Consider testing $H : \sigma \le \sigma_0$ versus $K : \sigma > \sigma_0$ Then $G = \{g : g(\underline{x}) = \underline{x} + c\underline{1}\}$ leaves H versus K invariant. By sufficiency, we can restrict attention to tests based on $\bar{X}$ and $S \equiv \sum (X_i - \bar{X})^2$. Let G^* denote the induced group $g^*(\bar{X}, S) = (\bar{X} + c, S)$. Thus S is a G^*MI by example 9. Now $S \sim \sigma^2\chi^2_{n-1}$, which has MLR in S . Thus by theorem 6.2.4.2, the UMP $G^* -$ invariant test of H versus K rejects H if $S > \sigma_0^2\chi^2_{n-1,\alpha}$. By theorem 6.5.6, Lehmann TSH (1986), page 301, it is also the UMP $G -$ invariant test.

3.4.4. Application 2. Two - sample t - test. Let $X_1, \cdots, X_m$ be iid $N(\mu, \sigma^2)$ and $Y_1, \cdots, Y_n$ be iid $N(\nu, \sigma^2)$, and consider testing $H : \nu \le \mu$ versus $K : \nu > \mu$. By sufficiency we can restrict attention to tests based on $(\bar{X}, \bar{Y}, S)$ with $S = \sum (X_i - \bar{X})^2 + \sum (Y_j - \bar{Y})^2$. Then the group

$G = \{g : g(\underline{x}) = a\underline{x} + b\underline{1}\}, a > 0, b \in R\}$ leaves H versus K invariant, and if G^* denotes the induced group

$$g^*(\bar{X}, \bar{Y}, S) = (a\bar{X} + b, a\bar{Y} + b, a^2 S),$$

then $T(\bar{X}, \bar{Y}, S) = (\bar{Y} - \bar{X})/\sqrt{S}$ is a $G^* -$ MI. Note

$$t \equiv \frac{\sqrt{\frac{mn}{N}}(\bar{Y} - \bar{X})}{\sqrt{\frac{S}{N-2}}} = \sqrt{\frac{mn}{N}(N-2)}\,T \sim t_{m+n-2}(\delta)$$

with $\delta \equiv \sqrt{\frac{mn}{N}}(\nu - \mu)/\sigma$, and that H versus K is equivalent to $H' : \delta \le 0$ versus $K' : \delta > 0$. Since the non-central $t -$ distributions have MLR, the UMP $G^* -$ invariant test of H versus K is the two-sample $t -$ test, "reject H if $t > t_{m+n-2,\alpha}$."

3.4.5. Application 3. Sampling inspection by variables. Let $Y, Y_1, \cdots, Y_n$ be iid $N(\mu, \sigma^2)$. Let $p \equiv P(Y \le y_0) \equiv P(\text{good})$ for some fixed number y_0. Consider testing $H : p \ge p_0$ versus $K : p < p_0$. Now

$$\begin{aligned} p &= P(Y \le y_0) = P\left(\frac{Y - y_0 - (\mu - y_0)}{\sigma} \le \frac{y_0 - \mu}{\sigma}\right) \\ &= P\left(\frac{X - \theta}{\sigma} \le -\frac{\theta}{\sigma}\right) \\ &\qquad \text{where } X_i \equiv Y_i - y_0 \sim N(\theta \equiv \mu - y_0, \sigma^2) \\ &= \Phi\left(-\frac{\theta}{\sigma}\right) = 1 - \Phi\left(\frac{\theta}{\sigma}\right), \end{aligned}$$

or $\theta/\sigma = \Phi^{-1}(1 - p)$. Thus, on the basis of $X_1, \cdots, X_n$ we wish to test $H : \theta/\sigma \le c_0 \equiv \Phi^{-1}(1 - p_0)$ versus $K : \theta/\sigma > c_0$. Now $\bar{X}, S = \sqrt{S^2}$ are sufficient. Also, H versus K is invariant under the group G of example 11 with $c > 0$; and a GMI in the space of the sufficient statistic is $T = \sqrt{n}\bar{X}/S$. Now $T \sim t_{n-1}(\delta)$ where $\delta \equiv \sqrt{n}\theta/\sigma$, and the family of distributions has MLR in T. Also H versus K is equivalent to $H' : \delta \le \delta_0 \equiv \sqrt{n}\Phi^{-1}(1 - p_0)$ versus $K' : \delta > \delta_0$. Thus the UMP $G -$ invariant level α test of H versus K rejects H if $T > t_{n-1,\alpha}(\delta_0)$.

3.4.6. Application 4: ANOVA/General Linear Model. (Many sub-examples!)

The canonical form of ANOVA is as follows:

$\Theta :\quad Z \sim N_n(\eta, \sigma^2 I), \quad \eta \in V_k \subset R^n$

where V_k is a subspace of R^n with dimension $k < n$, and $\eta_i = 0, \quad i = k+1, \cdots, n$.

Θ_0 : $Z \sim N_n(\eta, \sigma^2 I)$, $\eta \in V_{k-r} \subset V_k \subset R^n$

where V_{k-r} is a subspace of V_k R^n with dimension $r < k$, and $\eta_i = 0$, $i = 1, \cdots, r, k+1, \cdots, n$.

We let

$$G_1 \equiv \left\{ \begin{matrix} g_1 : g_1 z = (z_1, \cdots, z_r, z_{r+1} + \Delta_{r+1}, \cdots, z_k + \Delta_k, z_{k+1}, \cdots, z_n) \\ \text{with } \Delta_i \in R \} \end{matrix} \right\} ;$$

$$G_2 = \left\{ g_2 : \begin{matrix} g_2 z = (z_1^*, \cdots, z_r^*, z_{r+1}, \cdots, z_k, z_{k+1}, \cdots, z_n) \\ \text{where } (z_1^*, \cdots, z_r^*) \text{ is an orthog. trans. of } (z_1, \cdots, z_r) \end{matrix} \right\} ;$$

$$G_3 = \left\{ g_3 : \begin{matrix} g_3 z = (z_1, \cdots, z_r, z_{r+1}, \cdots, z_k, z_{k+1}^*, \cdots, z_n^*) \\ \text{where } (z_{k+1}^*, \cdots, z_n^*) \text{ is an orthog. trans. of } (z_{k+1}, \cdots, z_n) \end{matrix} \right\} ;$$

$$G_4 = \{ g_r : g_4 z = cz \text{ where } c \neq 0 \} ;$$

and, finally,

$$G \equiv G_4 \oplus G_3 \oplus G_2 \oplus G_1 \equiv \{ g_4 \circ g_3 \circ g_2 \circ g_1 : g_i \in G_i , i = 1, \cdots, 4 \} .$$

Then H versus K is invariant under G .

Now $T_1(z) = (z_1, \cdots, z_r, z_{k+1}, \cdots, z_n)$ is a G_1MI.

In the space of the G_1MI, a G_2MI is $T_2(\underline{z}) = (\sum_1^r z_i^2, z_{k+1}, \cdots, z_n)$.

In the space of the $G_2 \oplus G_1$ MI, a G_3MI is $T_3(\underline{z}) = (\sum_1^r z_i^2, \sum_{k+1}^n z_i^2)$.

In the space of the $G_3 \oplus G_2 \oplus G_1$ MI, a G_4 MI is $T(z) = ((n-k)/r)(\sum_1^r z_i^2 / \sum_{k+1}^n z_i^2)$.

Now $T(z)$ is a GMI; thus any $G-$ invariant test function for H versus K is a function of $T(z)$ by theorem 3.2.6.

Similarly,

$(\sigma^2, \eta_1, \cdots, \eta_r)$ is a $\bar{G}_1$ *MI* ;

$(\sigma^2, \sum_{i=1}^r \eta_i^2)$is a $\bar{G}_2 \oplus \bar{G}_1$ *MI* and a $\bar{G}_3 \oplus \bar{G}_2 \oplus \bar{G}_1$ *MI* ;

and

$$\delta^2 \equiv \frac{1}{2} \lambda^2 = \frac{\sum_{i=1}^r \eta_i^2}{2\sigma^2} \text{ is a GMI} .$$

Thus the distribution of any invariant test depends only on δ^2 .

Now $T \sim F_{r,n-k}(\delta^2)$, which has MLR in T. Also, H versus K is equivalent to $H' : \delta = 0$ versus $K' : \delta > 0$. Thus the UMP $G-$invariant test of H versus K rejects H when $T > F_{r,n-k,\alpha}$.

Reduction to Canonical Form

The above analysis has been developed for the linear model in canonical form. Now the question is: how do we reduce a model stated in a more usual way to the canonical form? Suppose that

$$\underline{X} \sim N_n(\underline{\xi}, \sigma^2 I)$$

where $\underline{\xi} \equiv E\underline{X} = A\underline{\theta} \in \mathbf{L}$, where A is a (known) $n \times k$ matrix of rank k, θ is a $k \times 1$ vector of (unknown) parameters, and $\mathbf{L}$ is the $k-$dimensional subspace of R^n spanned by the columns of the matrix A. Let B be a given $r \times k$ matrix, and consider testing

$$H : B\underline{\theta} = \underline{0} \qquad \text{or} \qquad \underline{\xi} \in \mathbf{L}_1$$

where $\mathbf{L}_1$ is a $(k-r)-$dimensional subspace of R_n.

To transform this form of the testing problem to canonical form, let T be an $n \times n$ orthogonal matrix with:

(i) the last $n-k$ rows of T are $\perp$ to $\mathbf{L}$; i.e. $\perp$ to the columns of A.

(ii) the rows $r+1, \cdots, k$ of T span $\mathbf{L}_1$.

Then set $\underline{Z} \equiv T\underline{X}$. We compute

$$\underline{\eta} \equiv E\underline{Z} = TA\underline{\theta} = T\underline{\xi},$$

and note that:

(a) $\eta_{k+1} = \cdots = \eta_n$ always by (i).

(b) $\eta_1 = \cdots = \eta_r = 0$ under H by (ii) since the first r rows of T are $\perp$ to $\mathbf{L}_1$.

Now we will re-express the $F-$statistic we have derived in terms of the X's:

$$S^2(\underline{\eta}) \equiv \sum_{i=1}^{n} (Z_i - \eta_i)^2 = \sum_{i=1}^{k} (Z_i - \eta_i)^2 + \sum_{i=k+1}^{n} Z_i^2$$

$$\geq \sum_{i=k+1}^{n} Z_i^2$$

by taking $\eta_i = \hat{\eta}_i = Z_i$, $i = 1, \cdots, k$. But since T is orthogonal and $\underline{Z} = T\underline{X}$,

(a) $$S^2(\underline{\eta}) = ||\underline{Z} - \underline{\eta}||^2 = (\underline{Z} - \underline{\eta})^T(\underline{Z} - \underline{\eta})$$

$$= (\underline{X} - \underline{\xi})^T(\underline{X} - \underline{\xi}) = \sum_{i=1}^{n} (X_i - \xi_i)^2$$

so that

(b) $$\min_{\underline{\xi} \in \mathbf{L}} \sum_{i=1}^{n} (X_i - \xi_i)^2 = \sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2 = \sum_{i=k+1}^{n} Z_i^2$$

where $\hat{\underline{\xi}}$ is the Least Squares (LS) estimator of $\underline{\xi}$ under $\xi = A\underline{\theta} \in \mathbf{L}$. Similarly, under $H : \underline{\xi} \in \mathbf{L}_1$ (or $\eta_1 = \cdots = \eta_r = 0$ in the canonical form),

$$S^2(\underline{\eta}) = \sum_{i=1}^{r} Z_i^2 + \sum_{i=r+1}^{k} (Z_i - \eta_i)^2 + \sum_{i=k+1}^{n} Z_i^2$$
$$\geq \sum_{i=1}^{r} Z_i^2 + \sum_{i=k+1}^{n} Z_i^2$$

by taking $\eta_i = \hat{\hat{\eta}}_i = Z_i$ for $i = r+1, \cdots, k$, and hence by (a)

(c) $$\min_{\underline{\xi} \in \mathbf{L}_1} \sum_{i=1}^{n} (X_i - \xi_i)^2 \equiv \sum_{i=1}^{n} (X_i - \hat{\hat{\xi}}_i)^2 = \sum_{i=1}^{r} Z_i^2 + \sum_{i=k+1}^{n} Z_i^2$$

where $\hat{\hat{\underline{\xi}}}$ is the least squares estimate of $\underline{\xi}$ under the hypothesis $\underline{\xi} \in \mathbf{L}_1$. Combining (b) and (c) yields

(d) $$\sum_{i=1}^{r} Z_i^2 = \sum_{i=1}^{n} (X_i - \hat{\hat{\xi}}_i)^2 - \sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2 ;$$

here $\mathbf{L}_1$ is a subspace of dimension $k - r$ contained in $\mathbf{L}$, which is a subspace of dimension k contained in R^n. Now

(e) $$\underline{X} - \underline{\xi} \perp \hat{\underline{\xi}} - \hat{\hat{\underline{\xi}}} \in \mathbf{L} .$$

Hence

$$||\underline{X} - \hat{\hat{\underline{\xi}}}||^2 = ||\underline{X} - \hat{\underline{\xi}}||^2 + ||\hat{\underline{\xi}} - \hat{\hat{\underline{\xi}}}||^2$$

by (e), and we have

$$\sum_{i=1}^{r} Z_i^2 = ||\hat{\underline{\xi}} - \hat{\hat{\underline{\xi}}}||^2 = \sum_{i=1}^{n} (\hat{\xi}_i - \hat{\hat{\xi}}_i)^2 ,$$

and the $F -$ statistic which yields the UMP $G -$ invariant test $H : \underline{\xi} \in \mathbf{L}_1$ versus $K : \underline{\xi} \notin \mathbf{L}_1$ may be written as

$$F = \frac{\sum_{i=1}^{n} (\hat{\xi}_i - \hat{\hat{\xi}}_i)^2/r}{\sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2/(n-k)} = \frac{[\sum_{i=1}^{n} (X_i - \hat{\hat{\xi}}_i)^2 - \sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2]/r}{\sum_{i=1}^{n} (X_i - \hat{\xi}_i)^2/(n-k)} .$$

To re-express the noncentrality parameter of the distribution of F under the alternative hypothesis in terms of $\underline{\xi}$ (instead of $\underline{\eta}$), let $\underline{\xi} \in \mathbf{L}$, and let $\underline{\xi}^0$ denote the projection of $\underline{\xi}$ onto $\mathbf{L}_1$: thus $\underline{\xi} = \underline{\xi}^0 + \underline{\xi} - \underline{\xi}^0$ where $\underline{\xi}^0 \in \mathbf{L}_1$ and $\underline{\xi} - \underline{\xi}^0 \perp \mathbf{L}_1$. Then

$$\delta^2 =$$

3.5. Rank tests

First we need to be able to compute probabilities for rank vectors. Our first job here is to develop a fundamental formula due to Hoeffding which allows us to do this.

Let $Z_1, \cdots, Z_N$ be independent with densities $f_1, \cdots, f_N$ respectively. Let $R_i \equiv$ rank of Z_i in $Z_1, \cdots, Z_N = \#\{j \le N : Z_j \le Z_i\}$. Thus

$$P(\underline{R} = \underline{r}) = \int_S \cdots \int f_1(z_1) \cdots f_N(z_N)\, dz_1 \cdots dz_N$$

where

$$\begin{aligned} S &\equiv [\underline{z} : R_i(\underline{z}) = r_i, \, i = 1, \cdots, N] \\ &= [\underline{z} : z_{d_1} < \cdots < z_{d_N}] \end{aligned}$$

where $d = r^{-1}$, the inverse permutation, $r \circ d = r \circ r^{-1} = e$.
[Example: $N = 3$; $z = (10, 5, 8)$. Then $r = (3, 1, 2)$ and $d = (2, 3, 1)$.]
Hence, letting $z_{d_i} \equiv v_i$,

$$S = [V_1 < \cdots < V_N],$$

and

$$\begin{aligned} P(\underline{R} = \underline{r}) &= \int_{v_1 < \cdots < v_N} \cdots \int \frac{f_1(v_{r_1}) \cdots f_N(v_{r_N})}{N! h(v_{r_1}) \cdots h(v_{r_N})} N! h(v_1) \cdots h(v_N)\, d\underline{v} \\ &= \frac{1}{N!} E \frac{f_1(V_{(r_1)}) \cdots f_N(V_{(r_N)})}{h(V_{(r_1)}) \cdots h(V_{(r_N)})} \end{aligned}$$

where $V_{(1)} < \cdots < V_{(N)}$ are the order statistics of a sample of size N from h. This formula is one version of *Hoeffding's formula.*

Of course, sometimes direct calculation succeeds immediately. Here are two simple, but important, examples:

Example 3.5.1. Suppose that $F_i = F^{\Delta_i}$ with $\Delta_i > 0$, $i = 1, \cdots, N$ and F continuous. Then

$$\begin{aligned} P(\underline{R} = \underline{e}) &= P(X_1 \le \cdots \le X_N) \\ &= \int_{x_1 \le \cdots \le x_N} \cdots \int \prod_{i=1}^N dF^{\Delta_i}(x_i) \\ &= \int_{0 \le u_1 \le \cdots \le u_N \le 1} \cdots \int \prod_{i=1}^N \Delta_i\, u_i^{\Delta_i - 1}\, du_i \\ &= \int_{0 \le u_2 \le \cdots \le u_N \le 1} \cdots \int \prod_{i=3}^N \Delta_i\, u_i^{\Delta_i - 1}\, \Delta_2\, u_2^{\Delta_1 + \Delta_2 - 1}\, du_2 \cdots du_N \end{aligned}$$

$$= \cdots = \prod_{i=1}^{N} \frac{\Delta_i}{\sum_{j=1}^{i} \Delta_j} .$$

This yields any probability $P(\underline{R} = \underline{r})$, $r \in \Pi$, by relabeling:

$$P(\underline{R} = \underline{r}) \; = \; P(Z_{d_1} \leq \cdots \leq Z_{d_N}) \; = \; \prod_{i=1}^{N} \frac{\Delta_{d_i}}{\sum_{j=1}^{i} \Delta_{d_j}} .$$

Example 3.5.2. Similarly, suppose that $(1 - F_i) = (1 - F)^{\Delta_i}$ with $\Delta_i > 0$, $i = 1, \cdots, N$ and F continuous; This is equivalent to $\Lambda_i \equiv -\log(1 - F_i) = \Delta_i\{-\log(1 - F)\} = \Delta_i \Lambda$, the proportional hazards model. Then

$$\begin{aligned} P(\underline{R} \; = \; \underline{e}) \; &= \; P(X_1 \leq \cdots \leq X_N) \\ &= \int_{x_1 \leq \cdots \leq x_N} \cdots \int \prod_{i=1}^{N} d\{1 - (1 - F(x_i))^{\Delta_i}\} \\ &= \int_{0 \leq u_1 \leq \cdots \leq u_N \leq 1} \cdots \int \prod_{i=1}^{N} \Delta_i (1 - u_i)^{\Delta_i - 1} \, du_i \\ &= \cdots = \prod_{i=1}^{N} \frac{\Delta_i}{\sum_{j=i}^{N} \Delta_j} , \end{aligned}$$

by integrating first over u_N , then $u_{N-1}, \cdots, u_1$. Again, this yields any probability $P(\underline{R} = \underline{r})$, $r \in \Pi$, by relabeling:

$$P(\underline{R} = \underline{r}) \; = \; P(Z_{d_1} \leq \cdots \leq Z_{d_N}) \; = \; \prod_{i=1}^{N} \frac{\Delta_{d_i}}{\sum_{j=i}^{N} \Delta_{d_j}} .$$

Now suppose that $X_1, \cdots, X_m$ are iid F and $Y_1, \cdots, Y_n$ are iid G , $F, G \in \mathbf{F}_c$; and let $\mathbf{G}$ denote the group of all strictly increasing continuous transformations of the real line onto itself, example 3.3.5.

Proposition 3.5.3.

A. The two - sample problem of testing $H : F = G$ versus $K : F <_s G$, $F, G \in \mathbf{F}_c$, is invariant under $\mathbf{G}$.

B. The rank vector $\underline{R}$ is a $\mathbf{G}$ – MI.

C. $\psi(u) = G \circ F^{-1}(u)$ is a $\bar{\mathbf{G}}$ – MI.

D. The ordered Y ranks $Q_1 < \cdots < Q_n$ are sufficient for $\underline{R}$;

$Q_i \equiv N\mathbb{H}_N(\mathbb{G}_n^{-1}(i/n))\ ,\ i = 1, \cdots, n$.

E. Hoeffding's formula: Suppose that F and G have densities f and g respectively, and that $f(x) = 0$ implies $g(x) = 0$. Then

$$P(\underline{Q} = \underline{q}) = \frac{1}{\binom{N}{n}} E_f \left\{ \prod_{j=1}^{n} \frac{g(V_{(q_j)})}{f(V_{(q_j)})} \right\}$$

where $V_{(1)} < \cdots < V_{(N)}$ are the order statistics of a sample of size N from F . Furthermore, this probability may be rewritten as

$$P(\underline{Q} = \underline{q}) = \frac{1}{\binom{N}{n}} E\{ \prod_{j=1}^{n} \psi'(U_{(q_j)})$$

where $U_{(1)} < \cdots < U_{(N)}$ are the order statistics of a sample of N uniform$(0,1)$ random variables.

Proof. E. From Hoeffding's formula with $f_i = f$ for $i = 1, \cdots, m$, $f_i = g$ for $i = m+1, \cdots, N$, and $h = f$, we obtain

$$P(\underline{R} = \underline{r}) = \frac{1}{N!} E_F(\prod_{j=1}^{n} \frac{g}{f}(V_{(r_{m+j})})) = \frac{1}{N!} E_F(\prod_{j=1}^{n} \frac{g}{f}(V_{(q_j)})) .$$

Hence

$$\begin{aligned} P(Q = q) &= \sum_{r\,:\,q(r)\,=\,q} P(\underline{R} = \underline{r}) \\ &= \frac{1}{N!} E_F(\prod_{j=1}^{n} \frac{g}{f}(V_{(q_j)})) \sum_{r\,:\,q(r)\,=\,q} 1 \\ &= \frac{m!\,n!}{N!} E_F(\prod_{j=1}^{n} \frac{g}{f}(V_{(q_j)})) . \end{aligned}$$

To see the second formula, note that $\psi'(u) = (g/f)(F^{-1}(u))$, and that $(F^{-1}(U_{(1)}), \cdots, F^{-1}(U_{(N)})) \stackrel{d}{=} (V_{(1)}, \cdots, V_{(N)})$. □

Proposition 3.5.4. Suppose that F has an absolutely continuous density f for which $\int |f'(x)|\, dx < \infty$. Then the locally most powerful rank test of $H: F = G$ versus $K: G = F(\cdot - \theta)$ with $\theta > 0$ is of the form

$$\phi(\underline{q}) = \begin{cases} 1 & \text{if } S_N \equiv \sum_{j=1}^{n} E_F(-\frac{f'}{f}(V_{(q_j)})) > k_\alpha \\ \gamma & \text{if } S_N = k_\alpha \\ 0 & \text{if } S_N < k_\alpha \end{cases}$$

where $V_{(1)} < \cdots < V_{(N)}$ are the order statistics in a sample of size N from F.

Proof. For a rank test, $\phi = \phi(\underline{Q})$, we want to maximize the slope of the power function at $\theta = 0$: i.e. to maximize the slope at $\theta = 0$ of

$$\beta_\phi(\theta) = E_\theta \phi(\underline{Q}) = \sum_{\underline{q}} \phi(\underline{q}) \, P_\theta(\underline{Q} = \underline{q}) .$$

To do this we clearly want to find those $\underline{q}$ for which

$$\frac{d}{d\theta} P_\theta(\underline{Q} = \underline{q})|_{\theta = 0}$$

is maximum. But, by using proposition 1 and differentiation under the expectation (which can be justified by the assumption that $\int |f'(x)| \, dx < \infty$),

$$\begin{aligned}
\frac{d}{d\theta} P_\theta(\underline{Q} = \underline{q})|_{\theta = 0} &= \frac{d}{d\theta} \frac{1}{\binom{N}{n}} E_F\{ \prod_{i=1}^{n} \frac{f(V_{(q_j)} - \theta)}{f(V_{(q_j)})} \}|_{\theta = 0} \\
&= \frac{1}{\binom{N}{n}} E_F\{ \frac{d}{d\theta} \prod_{i=1}^{n} \frac{f(V_{(q_j)} - \theta)}{f(V_{(q_j)})} |_{\theta = 0} \} \\
&= \frac{1}{\binom{N}{n}} E_F\{ - \sum_{j=1}^{n} \frac{f'}{f} (V_{(q_j)})\}
\end{aligned}$$

since

$$\begin{aligned}
\frac{d}{d\theta} \prod_{j=1}^{n} \frac{f(x_j - \theta)}{f(x_j)} |_{\theta = 0} &= \sum_{k=1}^{n} \prod_{j=1, j \neq k}^{n} \frac{f(x_j - \theta)}{f(x_j)} |_{\theta = 0} \{ - \frac{f'(x_k - \theta)}{f(x_k)} |_{\theta = 0} \\
&= - \sum_{k=1}^{n} \frac{f'}{f} (x_k) .
\end{aligned}$$

□

Example 3.5.5. If F is $N(\mu, \sigma^2)$, then without loss (by the monotone transformation $g(X) = (X - \mu)/\sigma$, we may take $F = N(0,1)$. Then $-(f'/f)(x) = x$, so $E\{ - \frac{f'}{f} (V_{(i)}) = E(Z_{(i)})$ where $Z_{(1)} < \cdots < Z_{(N)}$ are the order statistics of N standard normal ($N(0,1)$) rv's, and $S_N = \sum_{j=1}^{n} E(Z_{(q_j)})$. Note that $E(Z_{(i)})$ may be approximated by $\Phi^{-1}(i/(N+1))$ or by $\Phi^{-1}((3i-1)/(3N+1))$.

Example 3.5.6. If F is logistic, $f(x) = e^{-x}/(1+e^{-x})^2$, then $f = F(1-F)$, and $-f'/f = 2F - 1$. Since $F(V_{(i)}) \stackrel{d}{=} U_{(i)}$ where $U_1, \cdots, U_N$ are uniform$(0,1)$ rv's, with $E\,U_{(i)} = i/(N+1)$, the LMPRT of

$F = G$ versus $G = F(\cdot - \theta)$ rejects h for large values of $S_N = \sum_{j=1}^n Q_j$; this is the *Wilcoxon* statistic.

Proposition 3.5.7. Suppose that F has an absolutely continuous density f for which $\int |xf'(x)|\,dx < \infty$. Then the LMPRT of $H : F = G$ versus $K : G = F(\cdot/\theta)$, $\theta > 1$ and F specified is of the form

$$\phi(\underline{q}) = \begin{cases} 1 & \text{if } S_N \equiv \sum_{j=1}^{n} a_N(q_j) > k_\alpha \\ \gamma & \text{if } S_N = k_\alpha \\ 0 & \text{if } S_N < k_\alpha \end{cases}$$

where

$$a_N(i) \equiv E_F\{-1 - V_{(i)}\frac{f'}{f}(V_{(i)})\}$$

and $V_{(1)} < \cdots < V_{(N)}$ are the order statistics of a sample of size N from F.

Example 3.5.8. If $f(x) = e^{-x}1_{[0,\infty)}(x)$, then $(f'/f)(x) = -1$ and hence $a_N(i) = E_F\{-1 + V_{(i)}\} = E_F(V_{(i)}) - 1$ where $V_{(i)}$ are the order statistics of a sample of size N from F. But $V_{(i)} \stackrel{d}{=} \sum_{j=1}^{i}(Z_j/(N-j+1))$ where Z_j are iid exp(1), and hence

$$\begin{aligned} E(V_{(i)}) &= \sum_{j=1}^{i}\frac{1}{N-j+1} = \sum_{k=N-i+1}^{N}\frac{1}{k} \\ &= E\{-\log(1-U_{(i)})\} \end{aligned}$$

since $F^{-1}(t) = -\log(1-t)$. These are the Savage scores for testing exponential scale change; the approximate scores are

$$a_N(i) = -\log(1 - \frac{i}{N+1}), \quad i = 1, \cdots, N,$$

and the resulting test is sometimes called the "log - rank" test. It is derived nowadays in biostatistics via different considerations which allow for the introduction of censoring, and rewritten in a martingale framework.

[Recall that $\sum_{k=1}^N k^{-1} - \log N \to \gamma = .5772\cdots$, Euler's constant, as $N \to \infty$, so

$$\sum_{k=N-i+1}^{N}\frac{1}{k} = \sum_{k=1}^{N}\frac{1}{k} - \sum_{k=1}^{N-i}\frac{1}{k}$$

$$= \sum_{k=1}^{N} \frac{1}{k} - \log N - (\sum_{k=1}^{N-i} \frac{1}{k} - \log(N-i)) - \log(1 - \frac{i}{N})$$

$$\doteq -\log(1 - \frac{i}{N})$$

for large N .]

Note that when F is exponential(1) , then

$$(1 - G(x)) = 1 - F(\frac{x}{\theta}) = \exp(-\frac{x}{\theta}) = (1 - F(x))^{1/\theta} ,$$

or, $\Lambda_G = (1/\theta)\Lambda_F = \Delta\,\Lambda_F$ with $\Delta = 1/\theta$. Hence

$$\psi(u) = 1 - (1-u)^{1/\theta} = 1 - (1-u)^{\Delta} ,$$

and $\psi'(u) = \Delta(1-u)^{\Delta-1}$. Since the distribution of the ranks is the same for all (F, G) pairs with the same ψ , in fact the Savage test is the locally most powerful rank test of $H: F = G$ versus the Lehmann alternative $K: (1-G) = (1-F)^{\Delta}$, $\Delta < 1$.

Example 3.5.9. If F is $N(0, \sigma^2)$, the LMPRT of $F = G$ versus $K: G = F(\cdot/\theta)$, $\theta > 1$, rejects for large values of $S_N \equiv \sum_{j=1}^{n} a_N(Q_j)$ where $a_N(i) \equiv E(Z_{(i)}^2)$ where $Z_{(1)} < \cdots < Z_{(N)}$ is an ordered sample from $N(0,1)$. The approximate scores are $(\Phi^{-1}(i/(N+1)))^2$.

Remark. Note that any rank statistic of the form S_N can be rewritten in terms of empirical distributions as follows:

$$S_N = \sum_{i=1}^{n} a_N(i) = \sum_{j=1}^{n} a_N(R_{m+j}) = \sum_{i=1}^{N} a_N(i)\, Z_{Ni}$$

where $Z_{Ni} = 0$ or 1 according as the i th largest of the combined sample is an X or Y . Let $\mathbb{H}_N(x)$ = empirical df of the combined sample. Then $\mathbb{H}_N(i/N)$ = ith largest of the combined sample, $n\mathbb{G}_n(\mathbb{H}_N^{-1}(i/N))$ = the number of Y_i 's $\le \mathbb{H}_N^{-1}(i/N)$ and $Z_{Ni} = \Delta\{n\mathbb{G}_n(\mathbb{H}_N^{-1})\}(i/N)$. Therefore we can write

$$S_N = \sum_{i=1}^{N} a_N(i)\, Z_{Ni}$$

$$= \sum_{i=1}^{N} a_{Ni}\, \Delta\{n\mathbb{G}_n(\mathbb{H}_N^{-1})\}(i/N)$$

$$= n \int_0^1 \phi_N(u)\, d\mathbb{G}_n(\mathbb{H}_N^{-1}(u))$$

where $\phi_N(u)$ is the function defined by

$$\phi_N(u) \;=\; a_N(i) \quad \text{if} \quad \frac{i-1}{N} < u \le \frac{i}{N} \;,\; i = 1\,,\cdots,N \;.$$

If $\phi_N \to \phi$ and $\Lambda_N = m/N \to \lambda$, then it is often true that under alternatives $F \neq G$

$$\frac{1}{n}\, S_N \;=\; \int_0^1 \phi_N(u)\, d\mathbb{G}_n \circ \mathbb{H}_N^{-1}(u) \;\to_{a.s.}\; \int_0^1 \phi(u)\, dG \circ H^{-1}(u)$$

where $H = \lambda F + (1-\lambda) G$.

4. Efficiency of Tests

4.1. The Power of Two Tests

4.1.1 The power of the one-sample t-test: Let $X_1, \cdots, X_n$ be i.i.d. (θ, σ^2) We wish to test $H: \theta \le \theta_0$ versus $K: \theta > \theta_0$. The classical test of H versus K rejects H when $t_n \equiv \sqrt{n}(\bar{X} - \theta_0)/S > t_{n-1,\alpha}$.

(i) This test asymptotically correct level of significance (assuming $E(X^2) < \infty$ as we have) since

$$P_{\theta_0}(t_n > t_{n-1,\alpha}) \to P(N(0,1) > z_\alpha) = \alpha .$$

(ii) This test is consistent since, when a fixed $\theta > \theta_0$ is true

$$\begin{aligned} t_n &= \frac{\sqrt{n}(\bar{X} - \theta)}{S} + \frac{\sqrt{n}(\theta - \theta_0)}{S} \\ &\to_d N(0,1) + \infty \end{aligned}$$

and $t_{n-1,\alpha} \to z_\alpha$ so that $P_\theta(t_n > t_{n-1,\alpha}) \to 1$.

(iii) If $X_1, \cdots, X_n$ are i.i.d. $(\theta_n, \sigma^2) \equiv (\theta_0 + n^{-1/2}c_n, \sigma^2)$ where $c_n \to c$, then

$$\begin{aligned} t_n &= \frac{\sqrt{n}(\bar{X}_n - \theta_n)}{S} + \frac{c_n}{S} \\ &\to_d N(0,1) + \frac{c}{\sigma} = N(\frac{c}{\sigma}, 1) . \end{aligned}$$

Let $\beta_n^t(\theta)$ denote the power of the $t-$test based on $X_1, \cdots, X_n$ against the alternative θ. Then

$$\begin{aligned} \beta_n^t(\theta_n) &= \beta_n^t(\theta_0 + n^{-1/2}c_n) \\ &= P_{\theta_0 + c_n/\sqrt{n}}(t_n > t_{n-1,\alpha}) \to P(N(c/\sigma, 1) > z_\alpha) . \end{aligned} \tag{4.1}$$

4.1. The power of the Sign-test: Let $X_1, \cdots, X_n$ be i.i.d. with d.f. $F = F_0(\cdot - \theta)$ where F_0 has a unique median 0 (so that $F_0(0) = 1/2$. We wish to test $H: \theta \le \theta_0$ versus $K: \theta > \theta_0$. Let $Y_i \equiv 1_{[\theta_0,\infty)}(X_i)$, $i = 1, \cdots, n$. The *sign test* of H versus K rejects H when $S_n \equiv \sqrt{n}(\bar{Y} - 1/2)$ exceeds the upper α percentage point $s_{n,\alpha}$ of its distribution when θ_0 is true.

(i) When θ_0 is true, Y_i is Bernoulli(1/2) so that $S_n \to_d N(0, 1/4)$. Since the exact distribution of $\sum_{i=1}^n Y_i$ is Binomial$(n, 1/2)$ for all df's F as above, the test has exact level of significance α for all such F.

(ii) This test is consistent, since when a θ exceeding θ_0 is true

$$\begin{aligned} S_n &= \sqrt{n}(\bar{Y} - P_\theta(X > \theta_0)) + \sqrt{n}(P_\theta(X > \theta_0) - 1/2) \\ &\to_d N(0, p(1-p)) + \infty \end{aligned}$$

with $p \equiv 1 - F(\theta_0 - \theta) > 1/2$ so that $P_\theta(S_n > s_{n,\alpha}) \to 1$.

(iii) If $X_1, \cdots, X_n$ are i.i.d. $F = F_0(\cdot - (\theta_0 + n^{-1/2} d_n)$ where $d_n \to d$ as $n \to \infty$ and where we now assume that F_0 has a strictly positive derivative f_0 at 0. Then, using $F_0(0) = 1/2$, we have

$$\begin{aligned} S_n &= \sqrt{n}(\overline{Y}_n - P_{\theta_0 + n^{-1/2} d_n}(X > \theta_0)) + \sqrt{n}(P_{\theta_0 + n^{-1/2} d_n}(X > \theta_0) - 1/2) \\ &= \frac{1}{\sqrt{n}}(Binomial(n, 1 - F_0(-d_n/\sqrt{n})) - n(1 - F_0(-d_n/\sqrt{n}))) \\ &\quad + \sqrt{n}(F_0(0) - F_0(-d_n/\sqrt{n})) \\ &\to_d N(0, 1/4) + d\, f_0(0) = N(d\, f_0(0), 1/4). \end{aligned}$$

Thus the power of the sign test $\beta_n^s(\theta)$ satisfies

$$\beta_n^s(\theta_0 + n^{-1/2} d_n) \to P(\, N(d\, f_0(0), 1/4) > z_\alpha/2\,) = P(\, N(2d\, f_0(0), 1) > z_\alpha\,). \tag{4.2}$$

4.1.3. Problem. Use Liapunov's central limit theorem to show that

$$\frac{1}{\sqrt{n}}(Binomial(n, p_n) - np_n) \to_d N(0, p(1-p))$$

provided $p_n \to p$ as $n \to \infty$.

4.2. Pitman Efficiency.

4.2.1. Definition. *Pitman efficiency* is defined to be the limiting ratio of the sample sizes that produce equal asymptotic power against the same sequence of alternatives.

Now equal asymptotic power β in (4.1) and (4.2) requires that

$$\frac{c}{\sigma} = 2\, d\, f_0(0). \tag{4.3}$$

If the $t -$ test is based on N_t observations and the sign test based on N_s observations, then equal alternatives in (1) and (2) requires that

$$c_{N_t}/\sqrt{N_t} = d_{N_s}/\sqrt{N_s}. \tag{4.4}$$

Thus the Pitman efficiency $e_{s,t}$ of the sign test with respect to the $t -$ test is just the limiting value of N_t/N_s subject to (4.3) and (4.4). Thus

$$\frac{N_t}{N_s} = \frac{c_{N_t}^2}{d_{N_s}^2} \to (\frac{c}{d})^2 = 4\sigma^2 f_0^2(0) = e_{s,t}.$$

4.2.2. Problem. Evaluate $e_{s,t} = 4\sigma^2 f_0^2(0)$ in case:

(i) f_0 is uniform$(-a, a)$.

(ii) f_0 is Normal$(0, a^2)$.

(iii) f_0 is Logistic$(0, a)$: $f_0(x) = a^{-1}e^{-x/a}/[1+e^{-x/a}]^2$ with variance $= \pi^2 a^2/3$.

(iv) f_0 is t with k degrees of freedom.

(v) f_0 is double - exponential(a); $f_0(x) = (2a)^{-1}\exp(-a|x|)$.

4.2.3. A General Calculation We now consider the problem more generally. Suppose that $X_1, \cdots, X_N$ have a joint distribution P_θ where θ is a real - valued parameter. We wish to test $H: \theta \le \theta_0$ versus $K: \theta > \theta_0$. Suppose that the T_1 test and the T_2 test are both consistent tests of H versus K ; and that the T_i test rejects H if the statistic $T_{N,i}$ exceeds the upper α percent point of its distributon when $\theta = \theta_0$. Since both tests are consistent, it useless to compare their limiting power on against fixed alternatives; hence we will compare their power on a sequence of alternatives that approach θ_0 from above at the rate $1/\sqrt{N}$.

Suppose that for each $c > 0$ the statistics T_{Ni} satisfy

$$P_{\theta_0 + c_N/\sqrt{N}}(T_{N,i} \le x) \to P(N(c\mu_i, \sigma_i^2) \le \sigma_i x) = P(N(c\mu_i/\sigma_i, 1) \le x) \tag{4.5}$$

for all x as $N \to \infty$ for any sequences of c_N's converging to c. Let the $T_1 -$ test (the $T_2 -$ test) use N_1 (use N_2) observations against the sequence of alternatives $c_{N_1}/\sqrt{N_1}$ (the sequence of alternatives $c_{N_2}/\sqrt{N_2}$). where $c_{N_1} \to c_1$ (where $c_{N_2} \to c_2$). Equal asymptotic power requires

$$\frac{c_1\mu_1}{\sigma_1} = \frac{c_2\mu_2}{\sigma_2},$$

ans equal alternatives requires

$$\frac{c_{N_1}}{\sqrt{N_1}} = \frac{c_{N_2}}{\sqrt{N_2}};$$

solving these simultaneously leads to

$$\frac{N_2}{N_1} = \frac{c_{N_2}^2}{c_{N_1}^2} \to \frac{(\mu_1/\sigma_1)^2}{(\mu_2/\sigma_2)^2} = e_{1,2}. \tag{4.6}$$

Note that the efficiency $e_{1,2}$ is independent of the common level of significance α of the tests, of the particular value of the asymptotic power β, and of the particular sequences that converge to the values of c_1 and c_2 that are specified by the choice of β. Since so much is summarized in a single number, the procedure is bound to have some shortcomings; however it can be extremely useful and informative.

The quantity $\epsilon_i \equiv (\mu_i/\sigma_i)^2$ is called the *efficacy* of the $T_i -$ test, and hence the efficiency $e_{1,2}$ is the ratio of the efficacies.

4.2.4. Problem. Define your idea of what the exact small sample efficiency $e_{s,t}(\alpha, \beta, n)$ of the sign test with respect to the $t -$ test should be. Compute some values of it in in case $X_1, \cdots, X_n$ are normal, and compare these values with the asymptotic value $e_{s,t} = 2/\pi \doteq .6366\ldots$ that was obtained in problem 4.2.2.

4.2.5. Problem. Now redefine Pitman efficiency to be the ratio of the squared distances from the alternative to the hypothesized value θ_0 that produce equal asymptotic power as equal sample sizes approach infinity. Show that you get the same answer as before.

Note that if T_1 and T_2 are estimating the same thing (that is $\mu_1 = \mu_2$), then $e_{1,2}$ is just the ratio of the limiting variances.

Also note that the typical test of $H : \theta \le \theta_0$ versus $K : \theta > \theta_0$ is of the form: reject H if

$$\frac{\sqrt{n}(T - E_{\theta_0}(T))}{Var_{\theta_0}(\sqrt{n}T)} > constant_\alpha \to z_\alpha .$$

Thus when $\theta_0 + c/\sqrt{n}$ is true, intuitively we have (letting $m(\theta) \equiv E_\theta(T)$ and $\sigma_0^2 \equiv Var_{\theta_0}[\sqrt{n}T]$,

$$\frac{\sqrt{n}(T - E_{\theta_0}(T))}{\sqrt{Var_{\theta_0}(\sqrt{n}T)}} = \frac{\sqrt{Var_{\theta_0 + c/\sqrt{n}}(T)}}{\sqrt{Var_{\theta_0}(T)}} \frac{\sqrt{n}(T - E_{\theta_0 + c/\sqrt{n}}(T))}{\sqrt{Var_{\theta_0 + c/\sqrt{n}}(\sqrt{n}T)}} + \frac{\sqrt{n}(m(\theta_0 + c/\sqrt{n}) - m(\theta_0))}{\sqrt{Var_{\theta_0}(\sqrt{n}T)}}$$

$$\to_d 1 \cdot N(0,1) + \frac{cm'(\theta_0)}{\sigma_0} = N(cm'(\theta_0)/\sigma_0 , 1) .$$

Thus we expect $(m'(\theta_0)/\sigma_0)^2$ to be the efficacy.

4.2.6. Problem. Now consider testing $H : \theta = \theta_0$ versus $K : \theta \ne \theta_0$ on the basis of a two - sided test based on either the T_1 or the T_2 statistics considered previously. Show that the same formula for Pitman efficiency is appropriate for the two - sided test also.

4.2.7. Problem. Again consider testing $H : \theta = \theta_0$ versus $K : \theta \ne \theta_0$; but supppose now that that

$$T_{n,i} \to_d \chi_k^2(c^2\delta_i^2) \quad \text{as} \quad n \to \infty$$

under any sequence of alternatives $\theta_0 + c_n/\sqrt{n}$ having $c_n \to c > 0$ as $n \to \infty$. Here k is a fixed integer, and the limiting random variable has a noncentral chi - square distribution. Show that the Pitman efficiency criteria leads to $\epsilon_{1,2} = \delta_1^2/\delta_2^2$.

4.3. Some two-sample tests.

4.3.1. The two-sample t - test: Let $X_1, \cdots, X_m$ be independent samples from F and $G = F(\cdot - \theta)$ respectively. The classical test of $H: \theta \le 0$ versus $K: \theta > 0$ rejects H if

$$t \equiv \sqrt{\frac{mn}{N}}(\bar{Y} - \bar{X})/[\frac{m-1}{N-2} S_X^2 + \frac{n-1}{N-2} S_Y^2]^{1/2} > t_{m+n-2,\alpha} .$$

(This test can be shown to possess certain optimality properties when F is a normal distribution.) If F is any df having finite variance, then:

(i) When $\theta = 0$ we have $t \to_d N(0,1)$ provided $m \bigwedge n \to \infty$.

(ii) When $\theta > 0$ is true, then the test is consistent.

(iii) If $\lambda_N \equiv m/N \to \lambda \in (0,1)$ as $m \bigwedge n \to \infty$, then

$$P_{\theta = c/\sqrt{N}}(t > t_{n-2,\alpha}) \to P(N(c\sqrt{\lambda(1-\lambda)}/\sigma, 1) > z_\alpha) .$$

Thus the efficacy of the two-sample t - test is $\epsilon_t = \lambda(1-\lambda)/\sigma^2$.

4.3.2. The Mann-Whitney and Wilcoxon tests: Let $X_1, \cdots, X_m$ be i.i.d. F and let $Y_1, \cdots, Y_n$ be i.i.d. G where F and G are continuous df's, and consider testing $H: F = G$ versus $K: F <_s G$ (i.e. $G(x) \le F(x)$ with $<$ for some x).

The Wilcoxon test: reject H if $W_{m,n} \equiv \sum_{j=1}^{n} Q_j \equiv \sum_{j=1}^{n} R_{m+j}$ is "too big". Tables of the exact null distribution of $W_{m,n}$ for small m, n are available, so the level is exactly α . Moreover, if H is true

$$\frac{W_{m,n} - E_H(W_{m,n})}{\sqrt{Var_H(W_{m,n})}} = \frac{W_{m,n} - n(N+1)/2}{\sqrt{mn(N+1)/12}} \to_d N(0,1)$$

provided $m \bigwedge n \to \infty$; this follows from the WWNH permutational CLT since

$$\frac{N}{m \bigwedge n} \eta \equiv \frac{N}{m \bigwedge n} \frac{\max |a_i - \bar{a}|^2}{\sum (a_i - \bar{a})^2} = \frac{1}{m \bigwedge n} \frac{(N-1)^2/4}{(N^2-1)/12} \to 0$$

provided $m \bigwedge n \to \infty$.

The Mann - Whitney test: Let

$$U_{m,n} \equiv \frac{1}{m\,n} \sum_{i=1}^{m} \sum_{j=1}^{n} 1_{[X_i \le Y_j]} .$$

Mann and Whitney proposed to reject H if U is "too big". Since $m\,n\,U_{m,n} + n(n+1)/2 = W_{m,n}$, if H is true we have

$$\frac{U_{m,n} - 1/2}{\sqrt{(N+1)/12mn}} \to_d N(0,1) \quad \textit{provided} \quad m \bigwedge n \to \infty .$$

For arbitrary F and G

$$E\,U_{m,n} \;=\; E1_{[X \le Y]} \;=\; \int F\,dG$$

while for arbitrary continuous F and G

$$\begin{aligned} Var(\sqrt{mn}\,U_{m,n}) \;&=\; (n-1)\int (1-G)^2\,dF \;+\; (m-1)\int F^2\,dG \\ &\quad -\;(N-1)(\int F\,dG)^2 \;+\; \int F\,dG \\ &=\; (n-1)\,Var[1-G(X)] \;+\; (m-1)\,Var[F(Y)] \\ &\quad +\; \int F\,dG\,(1-\int F\,dG)\,. \end{aligned}$$

We now consider the local alternatives $Y \stackrel{d}{=} X + c/\sqrt{N}$, or $G \;=\; F(\cdot\;-\;c/\sqrt{N})$. We also suppose that $\lambda_N \;=\; m/N \;\to\; \lambda$. Then

$$\begin{aligned} \frac{U_{m,n}\;-\;1/2}{\sqrt{(N+1)/12m\,n}} \;&=\; \frac{U_{m,n}\;-\;\int F\,dG}{\sqrt{(N+1)/12m\,n}} \;+\; \frac{\int F\,dG\;-\;1/2}{\sqrt{(N+1)/12m\,n}} \\ &\equiv\; Z_{m,n} \;+\; a_{m,n} \end{aligned}$$

where it seems intuitive that $Z_{m,n} \;\to_d\; N(0,1)$ as $N\to\infty$ and

$$\begin{aligned} a_{m,n} \;&=\; \sqrt{\frac{12mn}{N(N+1)}}\,\sqrt{N}\{\int F\,dG \;-\; \frac{1}{2}\} \\ &=\; \sqrt{\frac{12mn}{N(N+1)}}\,\sqrt{N}\{\int F(x)\,dF(x-c/\sqrt{N}) \;-\; \int F\,dF\} \\ &=\; \sqrt{\frac{12mn}{N(N+1)}}\,\sqrt{N}\int \{F(x-c/\sqrt{N})-F(x)\}\,dF(x) \\ &\to\; \sqrt{12\lambda(1-\lambda)}\,c\int f^2(x)\,dx \end{aligned}$$

assuming that F has density f with $\int f^2(x)\,dx \;<\;\infty$. Thus under $(F\,,G) \;=\; (F\,,F(\cdot - c/\sqrt{N}))$

$$\frac{U_{m,n}\;-\;1/2}{\sqrt{(N+1)/12mn}} \;\to_d\; N(0,1) \;+\; c\sqrt{12\lambda(1-\lambda)}\int f^2 \;=\; N(c\sqrt{12\lambda(1-\lambda)}\int f^2\,,1)\,.$$

Thus the efficacy of the $U-$test is

$$\epsilon_U \;=\; 12\,\lambda(1-\lambda)\,\{\int f^2\}^2\,.$$

4.3.3. Pitman efficiency of the U − test wrt to the t − test. Combining the efficacies of 4.3.1 and 4.3.2 yields

$$\epsilon_{U,t}(F) \;=\; \frac{12\,\lambda(1-\lambda)\{\int f^2\}^2}{\lambda(1-\lambda)/\sigma^2} \;=\; 12\sigma^2\{\int f^2\}^2 \,.$$

4.3.4. Proposition. $\epsilon_{U,t}(F) \;\ge\; 108/125 \;=\; .864\cdots$.

proof. First note that $\epsilon_{U,t}(F) = \epsilon_{U,t}(F((\,\cdot - a)/c))$. Thus if suffices to minimize $\int f^2(x)\,dx$ subject to the restrictions

$$\int x^2 f(x)\,dx \;=\; 1\,, \qquad \int x\,f(x)\,dx \;=\; 0\,, \qquad f \;\ge\; 0\,, \qquad \int f(x)\,dx \;=\; 1\,.$$

Consider minimizing

$$B(f) \;\equiv\; \int_{-\infty}^{\infty} \{f^2(x) \;+\; f(x)2b(x^2 - a^2)\}\,dx \qquad \text{with} \quad b \;>\; 0$$

subject to $f \;\ge\; 0$ and $\int f(x)\,dx = 0$. Now

$$\begin{aligned} f^2(x) \;+\; 2bf(x)(x^2 - a^{2)} \;&=\; f(x)\{f(x) \;+\; 2b(x^2 - a^2)\} \\ &=\; A\{A \;+\; 2b(x^2 - a^2)\} \;\;\ge\; 0 \qquad \text{for } |x| \ge a\,. \end{aligned}$$

Thus take $f(x) = 0$ for $|x| > a$ and minimize the integrand pointwise for $|x| \le a$. This yields $A \;\equiv\; f(x) \;=\; b(a^2 - x^2)$. Thus the minimizer $f_{a,b}(x) \;\equiv\; f(x) \;=\; b(a^2 - x^2)1_{[-a,a]}(x)$. Choosing a and b so that $\int x^2\,f(x)\,dx \;=\; 1$ and $\int f(x)\,dx = 1$ yields $a \;=\; \sqrt{5}$, $b = 3\sqrt{5}/100$, and hence $\int f^2(x)\,dx \;=\; 3\sqrt{5}/25$. Hence

$$\epsilon_{U,t}(F) \;\ge\; 12\{\int f_{a,b}^2(x)\,dx\}^2 \;=\; 12\cdot\frac{9\cdot 5}{625} \;=\; \frac{108}{125}\,.$$

4.3.5. Proof of asymptotic normality of $U_{m,n}$ under local alternatives. Suppose that:

$$\begin{array}{lll} X_{N1}\,,\cdots,X_{Nm} & \text{are} & i.i.d. \quad F_N \\ Y_{N1}\,,\cdots,Y_{Nn} & \text{are} & i.i.d. \quad G_N \end{array}$$

where F_N, G_N and H are continuous df's satisfying $||F_N - H|| \to 0$ and $||G_N - H|| \to 0$ as $N \to \infty$. Let $I\!F_m$ and $I\!G_n$ denote the empirical df's of $X_{N1}\,,\cdots,X_{Nm}$ and $Y_{N1}\,,\cdots,Y_{Nn}$ respectively. Now

$$U_{m,n} \;=\; \int I\!F_m\, dI\!G_n\,.$$

Consider

$$
\begin{aligned}
(\frac{mn}{N})^{1/2}(U_{m,n} - \frac{1}{2}) &= (\frac{mn}{N})^{1/2}(\int \mathbb{F}_m \, d\mathbb{G}_n - \int F_N \, dG_N + \int F_N \, dG_N - \frac{1}{2}) \\
&= (\frac{mn}{N})^{1/2}\{ \int (\mathbb{F}_m - F_N) \, d\mathbb{G}_n + \int F_N \, d(\mathbb{G}_n - G_N)\} \\
&\quad + (\frac{mn}{N})^{1/2}\{ \int F_N \, dG_N - \int G_N \, dG_N\} \\
&\stackrel{d}{=} (\frac{n}{N})^{1/2} \int \mathbb{U}_m(F_N) \, d\mathbb{G}_n - (\frac{m}{N})^{1/2} \int \mathbb{V}_n(G_N) \, dF_N \\
&\quad + (\frac{N+1}{12\,N})^{1/2} a_{m,n} \, ;
\end{aligned}
$$

here $\mathbb{U}_m$ is the empirical process of m i.i.d. Uniform$(0,1)$ rv's and $\mathbb{V}_n$ is the empirical process of n i.i.d. Uniform$(0,1)$ rv's independent of the random variables used to define $\mathbb{U}_m$. Thus for special constructions of $\mathbb{U}_m$ and $\mathbb{V}_n$ and independent Brownian bridge processes $\mathbb{U}$ and $\mathbb{V}$,

$$
\begin{aligned}
(\frac{mn}{N})^{1/2}(U_{m,n} - \frac{1}{2}) &\to \bar{\lambda}^{1/2} \int \mathbb{U}(H) \, dH - \lambda^{1/2} \int \mathbb{V}(H) \, dH + \frac{1}{\sqrt{12}} \, a \\
&= \int_0^1 \{\bar{\lambda}^{1/2} \mathbb{U}(t) - \lambda^{1/2} \mathbb{V}(t)\} \, dt + \frac{a}{\sqrt{12}} \\
&\stackrel{d}{=} \int_0^1 \mathbb{U}(t) \, dt + \frac{a}{\sqrt{12}} = \frac{1}{\sqrt{12}} N(a\,,1)
\end{aligned}
$$

since

$$
\sigma^2 = \int_0^1 \int_0^1 (s \wedge t - st) \, ds dt = \frac{1}{12} \, .
$$

Convergence of the first term above is justified by:

$$
\begin{aligned}
&| \int \mathbb{U}_m(F_N) \, d\mathbb{G}_n - \int \mathbb{U}(H) \, dH | \\
&\le | \int [\mathbb{U}_m(F_N) - \mathbb{U}(F_N)] \, d\mathbb{G}_n | \\
&\qquad + | \int (\mathbb{U}(F_N) - \mathbb{U}(H)) \, d\mathbb{G}_n | + | \int \mathbb{U}(H) \, d(\mathbb{G}_n - H) | \\
&\le ||\mathbb{U}_m - \mathbb{U}|| \int d\mathbb{G}_n \\
&\qquad + ||\mathbb{U}(F_N) - \mathbb{U}(H)|| \int d\mathbb{G}_n + | \int \mathbb{U}(H) \, d(\mathbb{G}_n - H) | \\
&\to 0 + 0 + 0 = 0
\end{aligned}
$$

where the convergence of the first term follows by the special (Skorokhod) construction of $\{\mathbb{U}_m \, , \mathbb{U}\}$ and $\int d\mathbb{G}_n = 1$; the convergence of the second term follows from $||F_N - H|| \to 0$ and uniform continuity of $\mathbb{U}$ for a.e. fixed ω; and convergence of the third term follows from Helly - Bray since $\mathbb{U}(H)$ is a bounded continuous functon a.s. and $\mathbb{G}_n \to_d H$ almost surely. To understand this last claim better, note that

$$\begin{aligned} ||\mathbb{G}_n - H|| &= ||\mathbb{G}_n - G_N + G_N - H|| \\ &\le n^{-1/2}||\mathbb{V}_n(G_N)|| + ||G_N - H|| \\ &\le n^{-1/2}||\mathbb{V}_n(G_N) - \mathbb{V}(G_N)|| + n^{-1/2}||\mathbb{V}|| + ||G_N - H|| \\ &\to 0 + 0 + 0 = 0 . \end{aligned}$$

4.3.6. Problem. Evaluate $\epsilon_{U,t}(F) = 12\sigma^2(\int f^2(x)\,dx)^2$ in case:

(i) f is Uniform$(-a, a)$.

(ii) f is Normal.

(iii) f is logistic.

(iv) f is t_k.

(v) f is double exponential.

4.3.7. Problem. General behavior of the centering constants for $U_{m,n}$: Suppose that

$$||\sqrt{N}(f_N^{1/2} - h^{1/2}) - \frac{1}{2}\alpha h^{1/2}||_2 \to 0$$

and

$$||\sqrt{N}(g_N^{1/2} - h^{1/2}) - \frac{1}{2}\beta h^{1/2}||_2 \to 0 .$$

Then

$$||\sqrt{N}(F_N - H) - \int_{-\infty}^{\cdot} \alpha\, dH||_\infty \to 0$$

and

$$||\sqrt{N}(G_N - H) - \int_{-\infty}^{\cdot} \beta\, dH||_\infty \to 0 .$$

Show that this implies that

$$a_{m,n} \to \sqrt{12\lambda\bar{\lambda}} \int (1 - H)(\alpha - \beta)\, dH(y) .$$

4.4. Pitman Efficiency via Le Cam's third lemma.

Often the limiting power and efficacy of a test can be easily derived via Le Cam's third lemma, lemma 3.3.14. Recall that the essence of that lemma is that the joint limiting distribution of a statistic and the local log-likelihood ratio under the null hypothesis determines the joint limiting distribution of the statistic and the local log-likelihood ratio under the sequence of local alternatives. Here we simply illustrate this approach with the examples considered in section 6.4.1.

4.4.1. The one-sample t-test again. Let $X_1, \cdots, X_n$ be i.i.d $F = F_0(\cdot - \theta)$ with $\theta = E\,X$. Consider testing $H : \theta \le \theta_0$ versus $K : \theta > \theta_0$ using $t_n \equiv \sqrt{n}(\bar{X} - \theta_0)/S$. Suppose that F_0 has an absolutely continuous density f_0 and that $I_0 \equiv \int (f_0'/f_0)^2 f_0\,dx < \infty$. Let $L_n \equiv \prod_{i=1}^n (f_n/f)(X_i)$ where $f_n(x) \equiv f_0(x - \theta_n)$ and $\theta_n = \theta_0 + c/\sqrt{n})$. Thus, with $\dot{\mathbf{l}}(x) \equiv -(f'/f)(x)$,

$$\log L_n = \frac{c}{\sqrt{n}} \sum_{i=1}^n \dot{\mathbf{l}}(X_i) - \frac{c^2}{2} I(f_0) + o_p(1),$$

so,under P_n with $p_n(\underline{x}) = \prod_{i=1}^n f(x_i)$,

$$\begin{pmatrix} t_n \\ \log L_n \end{pmatrix} \overset{d}{\underset{P_n}{\to}} N_2\left(\begin{pmatrix} 0 \\ -(c^2/2)I_0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{tL} \\ \sigma_{tL} & c^2 I_0 \end{pmatrix} \right)$$

where

$$\sigma_{tL} = c\,E_f \frac{(X - \theta_0)}{\sigma} \dot{\mathbf{l}}(X) = c\,E_{f_0} \frac{X}{\sigma} \{ -\frac{f'_0(X)}{f_0(X)} \}.$$

But

$$1 = \frac{\theta_n - \theta_0}{c/\sqrt{n}} = \frac{1}{c/\sqrt{n}} \{E_{f_n}(X) - E_f(X)\}$$
$$= E_f\{X\,(\frac{(f_n/f)(X) - 1}{c/\sqrt{n}}),$$

so $1 = E_f\{X\,(-f'/f)(X)\}$. Hence it follows from Le Cam's third lemma with $q_n(\underline{x}) = \prod_{i=1}^n f_n(x_i)$,

$$\begin{pmatrix} t_n \\ \log L_n \end{pmatrix} \overset{d}{\underset{Q_n}{\to}} N_2\left(\begin{pmatrix} c/\sigma \\ +(c^2/2)I_0 \end{pmatrix}, \begin{pmatrix} 1 & c/\sigma \\ c/\sigma & c^2 I_0 \end{pmatrix} \right).$$

Hence the efficacy of the $t -$ test is $\epsilon_t = 1/\sigma^2$.

4.4.2. The sign-test again. Now consider the sign statistic $S_n = \sqrt{n}(\bar{Y} - 1/2)$ where $Y_i \equiv 1_{(\theta_0, \infty)}(X_i)$ and L_n is as above. Then under P_n, with $p_n(\underline{x}) \equiv \prod_1^n f(x_i)$,

$$\begin{pmatrix} S_n \\ \log L_n \end{pmatrix} \overset{d}{\underset{P_n}{\to}} N_2\left(\begin{pmatrix} 0 \\ -(c^2/2)I_0 \end{pmatrix}, \begin{pmatrix} 1/4 & \sigma_{SL} \\ \sigma_{SL} & c^2 I_0 \end{pmatrix} \right)$$

where

$$\begin{aligned} \sigma_{SL} &\equiv cE_f 1_{(\theta_0,\infty)}(X)\dot{\mathbf{l}}(X) \\ &= -c\int_{\theta_0}^{\infty} \frac{f'}{f}(x)\, f(x)\, dx \;=\; c\, f(\theta_0)\, . \end{aligned}$$

Hence, by Le Cam's third lemma

$$\begin{pmatrix} S_n \\ \log L_n \end{pmatrix} \xrightarrow[Q_n]{d} N_2\left(\begin{pmatrix} c\, f(\theta_0) \\ -(c^2/2)I_0 \end{pmatrix}, \begin{pmatrix} 1/4 & cf(\theta_0) \\ cf(\theta_0) & c^2 I_0 \end{pmatrix} \right),$$

and it follows that the efficacy of the sign test is $\epsilon_S = 4f^2(\theta_0) = 4f_0^2(0)$. Combining the two efficacies ϵ_t and ϵ_S yields the Pitman efficiency of the sign test relative to the $t-$test, $e_{s,t} = 4\sigma^2 f_0^2(0)$.