Statistics 583, Problem Set 1 Solutions

Wellner; 4/4/2007

1. (From Wasserman, *All of Statistics*, page 171; also see Wasserman's pages 161-164 for a brief discussion of permutation tests.) In 1861, 10 essays appeared in the *New Orleans Daily Crescent*. They were signed "Quintus Curtius Snodgrass" and some people suspected they were actually written by Mark Twain. To investigate this, we will consider the proportion of three letter words found in an author's work. From eight Twain essays we have

   .225, .262, .217, .240, .230, .229, .235, .217

   From 10 Snodgrass essays we have:

   .209, .205, .196, .210, .202, .207, .224, .223, .220, .201

   (a) Perform a Wald test for equality of the means. Give a $p-$value and a 95% confidence interval for the difference of means. What conclusion do you reach?
   (b) Now use a permutation test to avoid the use of large - sample methods. What is your conclusion?

   Solution: (a) Labelling the Twain proportions as X's and the Snodgrass proportions as Y's, we find that $\overline{X}_m = .231875$, $\overline{Y}_n = .2097$, $S_X = .01456$, and $S_Y = .00966$. Assuming that $X_i \sim N(\mu, \sigma^2$ and $Y_j \sim N(\nu, \tau^2)$ with $\sigma \neq \tau$, the Wald statistic becomes

$$\begin{aligned} W_{m,n} &= \left\{ \frac{\sqrt{\frac{mn}{N}}(\overline{X}_m - \overline{Y}_n)}{\sqrt{(n/N)S_X^2 + (m/N)S_Y^2}} \right\}^2 \\ &= \left\{ \frac{\sqrt{\frac{8\cdot 10}{18}}(.231875 - .2097)}{\sqrt{(10/18)(.000212125) + (8/18)(.0000933444)}} \right\}^2 \\ &= 3.70355^2 = 13.7163 \end{aligned}$$

   and the (approximate) p-value is $P(\chi_1^2 > 13.7163) = .000213$. If we use Welch's approximate $t-$test (see e.g. Lehmann and Casella, TSH, page 447), then the degrees of freedom f becomes, with $R \equiv mS_X^2/(nS_Y^2)$

$$\frac{1}{f} = \left(\frac{R}{1+R}\right)^2 \frac{1}{m-1} + \frac{1}{(1+R)^2}\frac{1}{n-1} = 1/13.6148.$$

Thus the approximate p-value using Welch's approximation is $P(|t_{13.61}| \geq 13.7163) =$ .00265. A 95% confidence interval for $\mu - \nu$ based on normal theory is given by

$$\begin{aligned} &\overline{X}_m - \overline{Y} \pm z_{.025}\sqrt{S_X^2/m + S_Y^2/n} \\ &= .022175 \pm 1.95996\sqrt{.000212125/8 + .0000933444/10} \\ &= (0.0104397, 0.0339103) \end{aligned}$$

The conclusion based on either of these tests is to reject the null hypothesis: from this evidence we would conclude that the Snodgrass and Twain essays were written by different authors.

(b) If we do an exact permutation t-test using the statistic introduced in class (involving the assumption of equal variances in the alternative), there are $\binom{18}{8} =$ 43758 combinations to consider, and the observed value of the statistic (in the form $(\overline{X}_m - \overline{z})/\sigma_N$) is 2.78917. By my calculations the exact one-sided p-value is 0.000525618, and the exact two-sided p-value is 0.000777. In contrast, by drawing $10^5 = 100,000$ random permutations, the estimated p-values were 0.00058 and 0.00088 respectively.

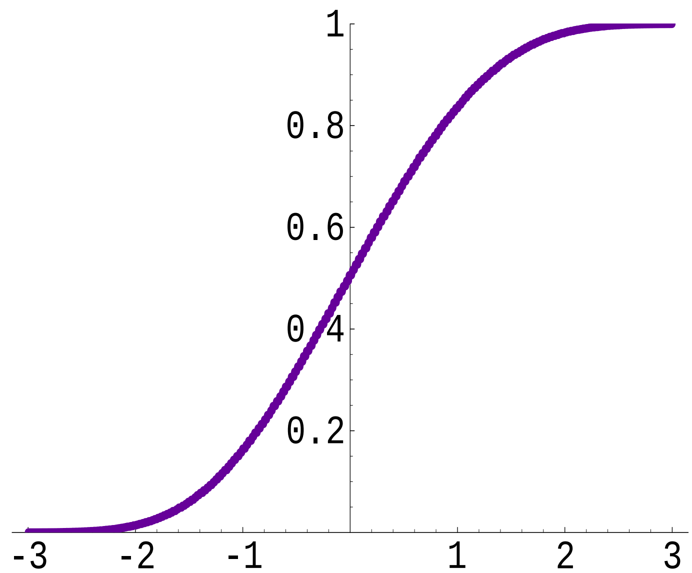

Figure 1: Exact permutation distribution, two-sample $t-$statistic $(\overline{X} - \overline{z})/\sigma_N$

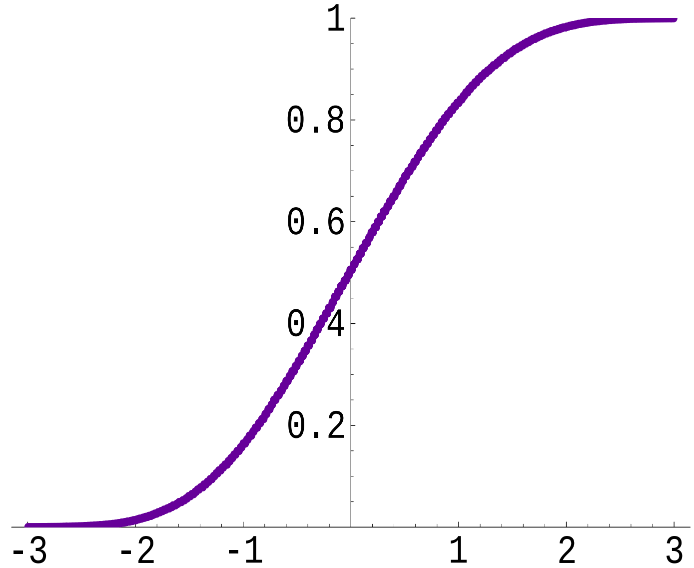

Figure 2: Approximate permutation distribution, two-sample $t-$statistic $(\overline{X} - \overline{z})/\sigma_N$, based on 10^5 random permutations

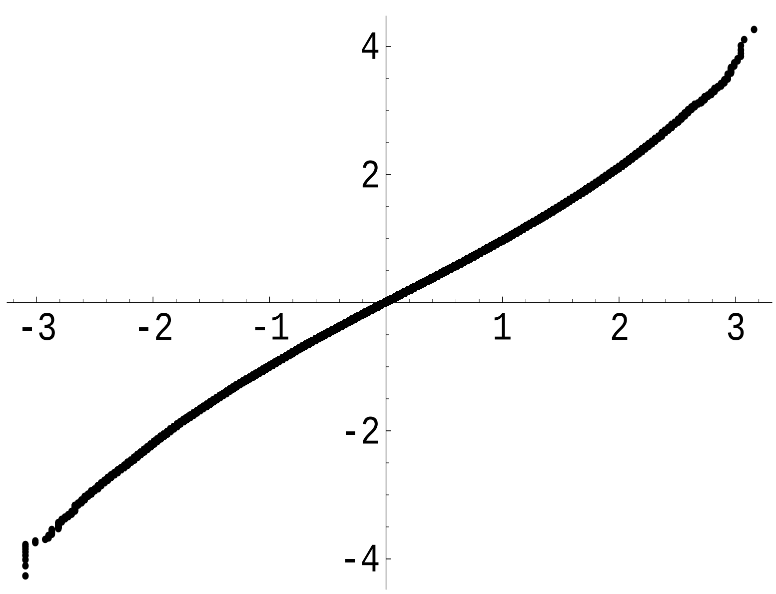

Figure 3: QQ-plot, Approximate permutation distribution, two-sample $t-$statistic $(\overline{X} - \overline{z})/\sigma_N$, based on 10^5 random permutations

I have not yet programmed the exact permutation test based on the Wald - type statistic used in part (a), but without squaring, which allows for the possibility of different variances. There are still $\binom{18}{8} = 43758$ combinations to consider, and the observed value of the test statistic is 3.70355. I have however, programmed the approximate permutation test based on sampling from the permutation distribution. By drawing $10^5 = 100,000$ random permutations, the estimated p-values I calculated are 0.00461 and 0.01064 respectively. It seems that the permutation distribution of the unequal variances version of the unsquared form of the Wald statistic is more nearly normal than that of the classical $t-$statistic. Note that while the two-sided permutation test still rejects at level $\alpha = .05$, this two-sided p-valued (0.01064)is not nearly as small as the estimated p-value of the permutation $t-$ test noted above (0.00088). We would continue to reject the null hypothesis at the level 0.05, but not at 0.01.

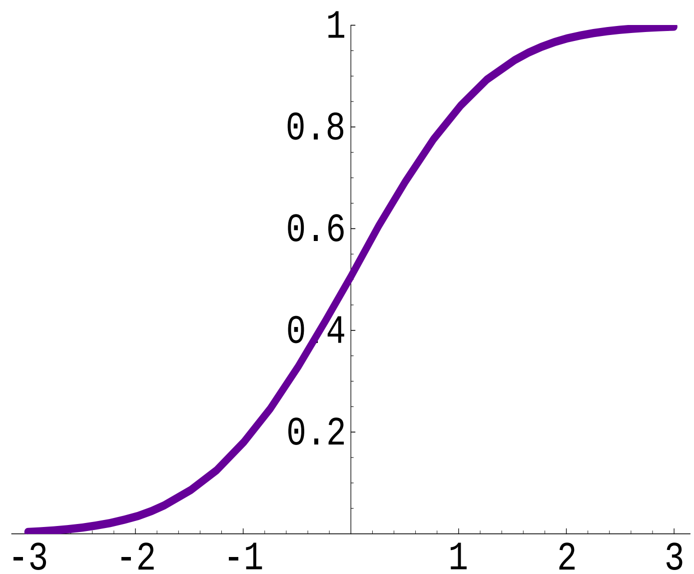

Figure 4: Approximate permutation distribution, one-sided Wald statistic $\overline{X} - \overline{Y})/\sqrt{S_X^2/m + S_Y^2/n}$ based on 10^5 random permutations

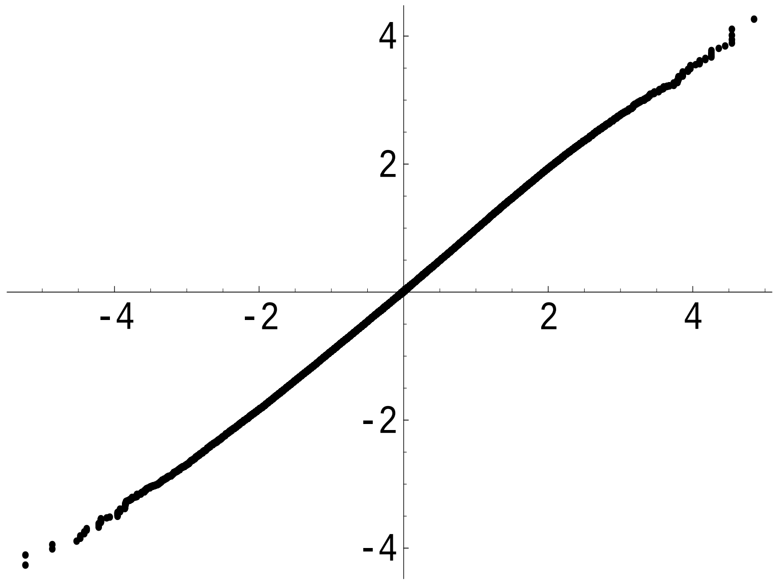

Figure 5: QQ-plot, Approximate permutation distribution, one-sided Wald statistic $\overline{X}-\overline{Y})/\sqrt{S_X^2/m+S_Y^2/n}$ based on 10^5 random permutations

2. For observations $\underline{X}=(X_1,\ldots,X_n)$, let $X_{(1)}\le\ldots\le X_{(n)}$ denote the *order statistics* of the X_i's ($X_{(i)}\equiv\mathbb{F}_n^{-1}(i/n)$, $i=1,\ldots,n$) and let $\underline{R}=(R_1,\ldots,R_n)$ denote the *ranks*; defined by $X_i=X_{(R_i)}$, $i=1,\ldots,n$ (if $X_i=X_j$ for some $i<j$, define the ranks by $R_i<R_j$ and $X_i=X_{(R_i)}$).
(a) Suppose that $X_1,\ldots,X_n$ are i.i.d. $F\in\mathcal{F}_{ac}$ (the absolutely continuous df's F on R) with density f. Show that the order statistics $\underline{X}_{(\cdot)}\equiv(X_{(1)},\ldots,X_{(n)})$ are independent of the ranks $\underline{R}$ and that the order statistics have joint density $\bar{p}$ given by

$$\bar{p}(\underline{x}_{(\cdot)})=n!\prod_{i=1}^n f(x_{(i)})\,,\qquad -\infty<x_{(1)}<\ldots<x_{(n)}<\infty$$

while

$$P(\underline{R}=\underline{r})=\frac{1}{n!}\,,\qquad \underline{r}\in\Pi\equiv\{\text{ all permutations of }\ \{1,\ldots,n\}\}\ .$$

(b) Show that the conclusion of (a) continues to hold for any joint distribution p of the $\underline{X}$ which is symmetric with respect to permutation of its coordinates: $p(\pi\underline{x})=p(\underline{x})$ for all $\underline{x}$ and $\pi\in\Pi$ where $\pi x\equiv(x_{\pi(1)},\ldots,x_{\pi(n)})$.
(c) If the joint distribution p of $\underline{X}$ is general (not permutation symmetric), show that the joint density $\bar{p}$ of the order statistics is given by

$$\bar{p}(\underline{x}_{(\cdot)})=\sum_{\pi\in\Pi}p(\pi\underline{x}_{(\cdot)})\ ,$$

and

$$P(\underline{R} = \underline{r} | \underline{X}_{(\cdot)} = \underline{x}_{(\cdot)}) = \frac{p(r\underline{x}_{(\cdot)})}{\bar{p}(\underline{x}_{(\cdot)})} \ .$$

Solution: I will prove C first; then A and B follow as corollaries:
C. Suppose that $\underline{X}$ has joint density p. Then for any set Borel set $A \subset \{\underline{x} \in \mathbb{R}^n : \ x_1 < x_2 < \cdots < x_n\}$

$$\begin{aligned} P(\underline{X}_{(\cdot)} \in A) &= \int_{[\underline{x}_{(\cdot)} \in A]} p(x_1, \ldots, x_n) \ dx_1 \ldots dx_n \\ &= \sum_{r \in \Pi} \int_{[R(\underline{x}) = r, \ \underline{x}_{(\cdot)} \in A]} p(x_1, \ldots, x_n) \ dx_1 \ldots dx_n \\ &= \sum_{r \in \Pi} \int_A p(x_{(r_1)}, \ldots, x_{(r_n)}) \ dx_{(1)} \ldots dx_{(n)} \\ &= \int_A \overline{p}(x_{(1)}, \ldots, x_{(n)}) \ dx_{(1)} \ldots dx_{(n)} \end{aligned}$$

where we have used the fact that the correspondence between $(x_1, ..., x_n)$ and $(x_{(1)}, \ldots, x_{(n)})$ is one-to-one and linear with Jacobian $= 1$ on each subset $[R = r]$, $r \in \Pi$. This proves that

$$\overline{p}(\underline{x}_{(\cdot)}) = \sum_{\pi \in \Pi} p(\pi \underline{x}_{(\cdot)}) \ .$$

Similarly,

$$\begin{aligned} P(R = r, \ \underline{X}_{(\cdot)} \in A) &= \int_{[R = r, \ \underline{x}_{(\cdot)} \in A]} p(x_1, \ldots, x_n) \ dx_1 \cdots dx_n \\ &= \int_A p(x_{(r_1)}, \ldots, x_{(r_n)}) \ dx_{(1)} \ldots dx_{(n)} \\ &= \int_A \frac{p(x_{(r_1)}, \ldots, x_{(r_n)})}{\overline{p}(x_{(1)}, \ldots, x_{(n)})} \overline{p}(x_{(1)}, \ldots, x_{(n)}) \ dx_{(1)} \ldots dx_{(n)} \end{aligned}$$

since $\overline{p}(x_{(1)}, \ldots, x_{(n)}) = 0$ implies $p(x_{r(1)}, \ldots, x_{r(n)}) = 0$ for each $r \in \Pi$. This implies that

$$P(\underline{R} = \underline{r} | \underline{X}_{(\cdot)} = \underline{x}_{(\cdot)}) = \frac{p(r\underline{x}_{(\cdot)})}{\overline{p}(\underline{x}_{(\cdot)})} \ .$$

B. When $p(\underline{x}) = p(\pi \underline{x})$ for all $\pi \in \Pi$, then

$$\overline{p}(\underline{x}_{(\cdot)}) = n! p(\underline{x}_{(\cdot)}),$$

and

$$P(\underline{R} = \underline{r} | \underline{X}_{(\cdot)}) = \frac{p(r\underline{x}_{(\cdot)})}{\overline{p}(\underline{x}_{(\cdot)})} = \frac{p(r\underline{x}_{(\cdot)})}{n!p(\underline{x}_{(\cdot)})} = \frac{1}{n!}.$$

Hence R is independent of $\underline{X}_{(\cdot)}$, and $P(R = r) = 1/n!$ for each $r \in \Pi$.
A. This follows easily from B since, in this case,

$$p(\underline{x}) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} f(x_{\pi(i)}) = p(\pi\underline{x}).$$

3. Let X and Y be independent exponential random variables with parameters λ and μ respectively: thus $P(X > x) = \exp(-\lambda x)$ and $P(Y > y) = \exp(-\mu y)$ for $x, y \geq 0$. Let $\theta \equiv \lambda/\mu$.
(a) Show that the problem of testing $H_0 : \theta \leq 1$ versus $H_1 : \theta > 1$ is invariant under the group G of transformations $g_c(x, y) = (cx, cy)$, $c > 0$, and find a UMP invariant test of size α.
(b) Show that the problem of testing $H_0' : \theta = 1$ versus $H_1' : \theta \neq 1$ is invariant *in addition* under the transformation $g(x, y) = (y, x)$, and find a UMP invariant test of size α.
(c) Find UMP invariant tests of the hypotheses in (a) and (b) when $X_1, \ldots, X_m$ are i.i.d. Exponential(λ) and $Y_1, \ldots, Y_n$ are i.i.d. Exponential(μ).

Solution: (a) Now X, Y have joint density

$$p_{\lambda,\mu}(x, y) = \lambda e^{-\lambda x} \mu e^{-\mu y} 1_{[0,\infty)}(x) 1_{[0,\infty)}(y)$$

so that if $c > 0$, $g_c(X, Y) = (cX, cY) \sim p_{\lambda/c,\mu/c}(x, y)$ and hence $\overline{g}(\lambda, \mu) = (\lambda/c, \mu/c)$. Note that $\delta(\lambda, \mu) = \lambda/\mu = (\lambda/c)/(\mu/c) = \delta(\overline{g}(\lambda, \mu))$, so that the hypotheses $H : \delta \leq 1$ and $K : \ \delta > 1$ are invariant. $T(X, Y) = X/Y$ and $\delta(\lambda, \mu) = \lambda/\mu$ are maximmal invariants on the sample space and parameter space respectively, and since $2\mu Y \sim \chi_2^2$ and $2\lambda X \sim \chi_2^2$ are independent

$$T(X, Y) = \frac{2\lambda X}{2\mu Y} \frac{\mu}{\lambda} \sim \delta^{-1} F_{2,2}.$$

Since this family of distributions has monotone (decreasing) likelihood ratio, it follows that the UMP $G-$invariant test of H versus K rejects H when $T < F_{2,2,\alpha}$ (where $P(F_{2,2} \leq F_{2,2,\alpha}) = \alpha$. or when $T^{-1} = Y/X > F_{2,2,1-\alpha} = (1-\alpha)/\alpha$.
(b) If $g(x, y) = (y, x)$ is also considered, then the transformation induced on the maximal invariant of A is given by

$$g(T) = T(g(X, Y)) = T(Y, X) = \frac{Y}{X} = \frac{1}{T(X, Y)}.$$

Claim: $T \vee T^{-1}$ is the maximal invariant with respect to this new group.
Proof. $T \vee T^{-1}$ is invariant since $T^{-1} \vee T = T \vee T^{-1}$; and $T \vee T^{-1}$ is maximal since $T \vee T^{-1} = T^* \vee T^{*-1}$ implies that either $T = T^*$ or $T = T^{*-1}$.

A corresponding maximal invariant on the parameter space is $\delta \vee \delta^{-1} = \frac{\lambda}{\mu} \vee \frac{\mu}{\lambda} \equiv \nu$, and the hypotheses $H : \delta = 1$ and $K : \delta \neq 1$ are clearly invariant. When expressed in terms of δ the hypotheses become $H : \delta = 1$ versus $K : \delta > 1$. It remains to show that the maximal invariant has monotone likelihood ratio (MLR).

By direct calculation, for $t > 1$,

$$1 - F_\eta(t) \quad = \quad P_{\lambda,\mu}(T \vee T^{-1} > t) = \frac{2 + \eta t}{1 + \eta t + t^2} \qquad (1)$$

with $\eta = \delta + 1/\delta$. Hence $T \vee T^{-1}$ has density

$$f_\eta(t) = \frac{\eta(t^2 + 1) + 4t}{(t^2 + \eta t + 1)^2} 1_{[1,\infty)}(t).$$

so that for $\eta_1 < \eta_2$

$$\frac{f_{\eta_2}}{f_{\eta_1}}(t) = \left(\frac{t^2 + \eta_1 t + 1}{t^2 + \eta_2 t + 1} \right)^2 \frac{\eta_2(t^2 + 1) + 4t}{\eta_1(t^2 + 1) + 4t} \equiv g(t)^2 h(t)$$

where

$$g'(t) = \frac{(\eta_2 - \eta_1)(t^2 - 1)}{(t^2 + \eta_2 t + 1)^2} \geq 0 \qquad \text{for} \quad t \geq 1$$

and

$$h'(t) = \frac{(\eta_2 - \eta_1)(4t^2 - 1)}{(t^2 + \eta_2 t + 1)^2} \geq 0 \qquad \text{for} \quad t \geq 1.$$

Hence the distribution of $T \vee T^{-1}$ has MLR and the UMP $G-$invariant test rejects H when $M \equiv T \vee T^{-1} > \frac{2-\alpha}{\alpha}$. Then, when $\delta = 1$, $\eta = 2$, and

$$P_{\eta=2}(M > (2 - \alpha)/\alpha)) = \alpha.$$

[Proof of (1): Since $T \sim \delta^{-1} F_{2,2}$ where $P(F_{2,2} \leq x) = x/(1 + x)$,

$$\begin{aligned} P(T \vee T^{-1} > t) &= P(T > t) + P(T < t^{-1}) = P(F_{2,2} > \delta t) + P(F_{2,2} < \delta/t) \\ &= \frac{1}{1 + \delta t} + \frac{\delta/t}{1 + \delta/t} \\ &= \frac{1/\delta}{1/\delta + t} + \frac{\delta}{\delta + t} \\ &= \frac{2 + (\delta + 1/\delta)t}{1 + (\delta + 1/\delta)t + t^2}.] \end{aligned}$$

(c) When $X_1, \ldots, X_m$ are i.i.d. exponential(λ) and $Y_1, \ldots, Y_n$ are i.i.d. exponential(μ) respectively, and $G = \{g : g(\underline{x}, \underline{y}) = (c\underline{x}, c\underline{y}), c > 0\}$, then we reduce by sufficiency first: $(\sum_1^m X_i, \sum_1^n Y_j)$ is sufficient for (λ, μ). Then $T = \sum_1^m X_i / \sum_1^n Y_j$ is a maximal invariant with respect to $G^* = \{g^* : g^*(x, y) = (cx, cy),\ c > 0\}$ acting on the space of the sufficient statistic. Moreover,

$$T = \frac{\mu}{\lambda}\frac{m}{n}\frac{2\lambda \sum_1^m X_i/(2m)}{2\mu \sum_1^n Y_j/(2n)} \sim \frac{1}{\delta n/m} F_{2m,2n}$$

where $\delta \equiv \delta(\lambda, \mu) = \lambda/\mu$ is a maximal invariant with respect to $\overline{G}$ on the parameter space, and the density of $F_{2m,2n}$ is given by

$$f_{F_{2m,2n}}(x) = c_{m,n} \frac{x^{m-1}}{(1 + (m/n)x)^{m+n}} 1_{(0,\infty)}(x).$$

Thus the density of T is given by

$$f_T(x, \delta) = \delta \tilde{c}_{m,n} \frac{(\delta x)^{m-1}}{(1 + \delta x)^{m+n}} 1_{(0,\infty)}(x)$$

which has monotone (decreasing) likelihood ratio in $M(x) = x$. Thus the UMP-G^* invariant test rejects H_0 when $((n/m)T)^{-1} = (\sum_1^n Y_j/2n)/(\sum_1^m X_i/2m) > F_{2n,2m,\alpha}$ where $P(F_{2n,2m} > F_{2n,2m,\alpha}) = \alpha$.

For testing $H_0' : \theta = 1$ versus $H_1' : \theta \neq 1$, we find, much as in (b), that $\tilde{T} \equiv T \vee T^{-1}$ is a maximal invariant with respect to the induced group $G^* = G_2^* \oplus G_1^*$ on the space of the sufficient statistic. By a calculation similar to that of part (b),

$$\begin{aligned}
1 - F(t) &= P(T \vee T^{-1} > t) = P(T > t) + P(T < 1/t) \\
&= P(\delta^{-1}(m/n)F_{2m,2n} > t) + P(\delta^{-1}(m/n)F_{2m,2n} < 1/t) \\
&= P((m/n)F_{2m,2n} > \delta t) + P((m/n)F_{2m,2n} < \delta/t)
\end{aligned}$$

and hence

$$\begin{aligned}
f_{T \vee T^{-1}}(t) &= \delta \tilde{c}_{m,n} \frac{(\delta t)^{m-1}}{(1 + \delta t)^{m+n}} + \frac{\delta}{t^2} \tilde{c}_{m,n} \frac{(\delta/t)^{m-1}}{(1 + \delta/t)^{m+n}} \\
&= \tilde{c}_{m,n} \frac{t^{m-1}(\delta + t)^m (1 + t/\delta)^n + t^{n-1}(1 + \delta t)^m (1/\delta + t)^n}{(t^2 + (\delta + 1/\delta)t + 1)^{m+n}}.
\end{aligned}$$

It is not yet clear to me that this density depends only on $\eta = \delta + 1/\delta$, so I may be making a mistake somewhere.

4. Suppose that $X \sim \text{Binomial}(m, p_1)$ and $Y \sim \text{Binomial}(n, p_2)$ are independent. In 582 lectures last quarter we derived the UMP unbiased (conditional) test of level α for testing $H : p_2 \le p_1$ versus $K : p_2 > p_1$. It involves rejecting H if $Y > c(t)$ relative to the conditional (hypergeometric) distribution of Y conditional on $T = X + Y = t$, or equivalently if

$$\frac{Y/t - (n/N)}{\sigma_N} > c'(t)$$

where $c'(T) \rightarrow_p z_\alpha \equiv \Phi^{-1}(1-\alpha)$ since

$$\frac{Y/t - (n/N)}{\sigma_N} \rightarrow_d Z \sim N(0,1)$$

conditionally on T if $0 < \liminf(n/N) \le \limsup(n/N) < 1$. Here $\sigma_N^2 = (1 - (t-1)/(N-1))(n/N)(1 - n/N)/t$.

(a) Show that

$$\frac{Y/t - (n/N)}{\sigma_N} = \frac{\sqrt{\frac{mn}{N}}(\frac{Y}{n} - \frac{X}{m})}{\sqrt{\frac{t}{N}(1 - \frac{t-1}{N-1})}}.$$

(b) Use the result of (a) to show that

$$\frac{Y/T - (n/N)}{\sigma_N} \rightarrow_d Z \sim N(0,1)$$

unconditionally under $p_1 = p_2 \equiv p$ (even if $a = \liminf(n/N) < \limsup(n/N) = b$ with possibly $a = 0$ or $b = 1$). [Hint: prove it first under the assumption that $n/N \rightarrow \lambda \in [0,1]$, and then show that the limit is the same even if n/N does not converge by considering subsequences.

(c) What is the (unconditional) limiting behavior of the test statistic $(Y/t - (n/N))/\sigma_N$ under local alternatives of the form $p_2 = p_{2,N} = p_1 + c/\sqrt{N}$ assuming that $n/N \rightarrow \lambda \in (0,1)$?

(d) What does the result of (c) imply about the limiting power of the test under these alternatives?

Solution: (a) Note that

$$\begin{aligned} \frac{Y}{T} - \frac{n}{N} &= \frac{1}{T}\left(Y - \frac{n}{N}T\right) = \frac{1}{T}\left(Y - \frac{n}{N}(X+Y)\right) \\ &= \frac{1}{T}\left(\frac{m}{N}Y - \frac{n}{N}X\right) = \frac{mn}{NT}\left(\frac{Y}{n} - \frac{X}{m}\right). \end{aligned}$$

Therefore

$$\frac{Y/t-(n/N)}{\sigma_N} = \frac{Y/t-n/N}{\sqrt{(1-\frac{n-1}{N-1})\frac{n}{N}\frac{m}{N}\frac{1}{t}}} = \frac{\sqrt{t}(Y/t-n/N)}{\sqrt{(1-\frac{n-1}{N-1})\frac{n}{N}\frac{m}{N}}}$$
$$= \frac{\frac{mn}{N}\left(\frac{Y}{n}-\frac{X}{m}\right)}{\sqrt{T(1-\frac{T-1}{N-1})\frac{mn}{N^2}}} = \frac{\sqrt{\frac{mn}{N}}\left(\frac{Y}{n}-\frac{X}{m}\right)}{\sqrt{\frac{T}{N}(1-\frac{T-1}{N-1})}}$$

as claimed.

(b) Now $(\sqrt{m}(X/m-p), \sqrt{n}(Y/n-p)) \to_d (Z_1, Z_2)$ where $Z_j \sim N(0, p(1-p))$, $j = 1, 2$ are independent. Thus from (a), under the assumption that $n/N \to \lambda \in [0,1]$,

$$V_N \equiv \frac{Y/T-(n/N)}{\sigma_N} = \frac{\sqrt{\frac{mn}{N}}\left(\frac{Y}{n}-\frac{X}{m}\right)}{\sqrt{\frac{T}{N}(1-\frac{T-1}{N-1})}}$$
$$= \frac{\sqrt{\frac{m}{N}}\sqrt{n}(Y/n-p)-\sqrt{\frac{n}{N}}\sqrt{m}(X/m-p)}{\sqrt{\frac{T}{N}(1-\frac{T-1}{N-1})}}$$
$$\to_d \frac{\sqrt{1-\lambda}Z_2-\sqrt{\lambda}Z_1}{\sqrt{p(1-p)}} \sim N(0,1)$$

where we used $T/N = (m/N)(X/m) + (n/N)(Y/n) \to_p (1-\lambda)p + \lambda p = p$.

If $\lambda_N \equiv n/N \not\to$, then since $\lambda_N \in [0,1]$, for any initial subsequence $\{\lambda_{N'}\}$, there exists a further convergent subsequence $\{\lambda_{N''}\}$; i.e. $\lambda_{N''} \to$ some $\lambda \in [0,1]$. By the same argument as above, for this subsequence $V_{N''} \to_d Z \sim N(0,1)$. Since the limiting distribution is the same for any such initial subsequence $\{V_{N'}\}$, we conclude that the full sequence $\{V_N\}$ satisfies $V_N \to_d Z \sim N(0,1)$ under $p_1 = p_2 = p$. (This argument is completely analogous to the following fact concerning real numbers: a sequence $\{x_n\}$ of real numbers satisfies $x_n \to x$ if and only if each subsequence $\{x_{n'}\}$ contains a further subsequence $\{x_{n''}\}$ such that $x_{n''} \to x$. See Billingsley (1968), *Convergence of Probability Measures*, theorem 2.3, page 16.)