Statistics 581, Problem Set 4 Solutions

Wellner; 10/23/2008

1. Suppose that $\underline{N}_n \sim$ Mult$_k(n, \underline{p})$ and $\widehat{p} = \underline{N}_n/n$. Suppose that the true $\underline{p}$ is $\underline{p}_n = \underline{p}_0 + n^{-1/2}\underline{c}$ where $\underline{1}^T\underline{c} = 0$. Use the Cramér - Wold device together with either the Liapunov or the Lindeberg-Feller CLT to show that

$$\underline{Z}_n = \left(\frac{N_{n,1} - np_{n,1}}{\sqrt{np_{0,1}}}, \ldots, \frac{N_{nk} - np_{n,k}}{\sqrt{np_{0,k}}} \right)$$

satisfies $\underline{Z}_n \rightarrow_d \underline{Z}$ where $\underline{Z} \sim N_k(0, I - \sqrt{p_0}\sqrt{p_0}^T)$. (It therefore follows, as outlined in class, that the chi-square statistic $Q_n \rightarrow_d \chi^2_{k-1}(\delta)$ with $\delta = \sum_{j=1}^k c_j^2/p_{0,j}$ under the local alternative $\underline{p}_n$.)

Solution: We argued heuristically in class that when the true $\underline{p} = \underline{p}_n = \underline{p}_0 + \underline{c}n^{-1/2}$, then

(1) $$\underline{Z}_n \equiv \text{diag}(1/\sqrt{\underline{p}_0})n^{1/2}(\widehat{\underline{p}} - \underline{p}_0) \rightarrow \underline{Z} + \underline{d} \sim N_k(\underline{d}, \Sigma)$$

where $\underline{d} = \text{diag}(1/\sqrt{\underline{p}_0})\underline{c}$ and $\Sigma = I - \sqrt{\underline{p}_0}\sqrt{\underline{p}_0}'$. To prove that (1) holds, we can use the Cramér-Wold device and the Liapunov CLT. Fix $\underline{a} \in R^k$. Then we want to show that

$$\underline{a}^T\sqrt{n}(\widehat{\underline{p}}_n - \underline{p}_n) \rightarrow_d N(0, \underline{a}^T(\text{diag}(\underline{p}_0) - \underline{p}_0\underline{p}_0^T)\underline{a}) .$$

But since $\underline{N}_n = \sum_{i=1}^n \underline{\Delta}_{ni}$ where $\underline{\Delta}_{ni} \sim$ Mult$_k(1, \underline{p}_n)$ are i.i.d. for each n, we can write

$$\begin{aligned} \underline{a}^T\sqrt{n}(\widehat{\underline{p}}_n - \underline{p}_n) &= \sum_{i=1}^n \sum_{j=1}^k a_j(\Delta_{ni,j} - p_{nj})/\sqrt{n} \\ &\equiv \sum_{i=1}^n X_{ni} \end{aligned}$$

where the X_{ni}'s have $\mu_{ni} = E(X_{ni}) = 0$,

$$\sigma^2_{ni} = Var(X_{ni}) = \underline{a}^T(\text{diag}(\underline{p}_n) - \underline{p}_n\underline{p}_n^T)\underline{a}/n$$

and

$$\gamma_{ni} = E|X_{ni}|^3 = n^{-3/2}\sum_{j'=1}^k \left\{ \left| a_{j'}(1 - p_{nj'}) + \sum_{j \neq j', j=1}^k a_j(0 - p_{nj}) \right|^3 \right\} p_{nj'} ,$$

so that

$$\sigma^2_n = \sum_1^n \sigma^2_{ni} = \underline{a}^T(\text{diag}(\underline{p}_n) - \underline{p}_n\underline{p}_n^T)\underline{a} \rightarrow \underline{a}^T\Sigma\underline{a}$$

while

$$\begin{aligned}
\gamma_n &= \sum_1^n \gamma_{ni} \\
&= n^{-1/2} \sum_{j'=1}^k \left\{ a_{j'}(1-p_{nj'}) + \sum_{j=1, j\neq j'}^k a_j(0-p_{nj})\Big|^3 \right\} p_{nj'} \\
&\to 0 \cdot M(\underline{a}, \underline{p}_0) = 0
\end{aligned}$$

where

$$M(\underline{a}, \underline{p}_0) = \sum_{j'=1}^k \left\{ a_{j'}(1-p_{0j'}) + \sum_{j=1, j\neq j'}^k a_j(0-p_{0j})\Big|^3 \right\} p_{0j'} \,.$$

hence it follows that $\gamma_n/\sigma_n^{3/2} \to 0$, and

$$\frac{\underline{a}^T \sqrt{n}(\widehat{\underline{p}}_n - \underline{p}_n)}{\sigma_n} = \frac{\sum_{i=1}^n X_{ni}}{\sigma_n} \to_d N(0,1)\,.$$

Since $\sigma_n^2 \to \underline{a}^T \Sigma \underline{a}$, this implies

$$\underline{a}^T \sqrt{n}(\widehat{\underline{p}}_n - \underline{p}_n) \to_d N(0, \underline{a}^T \Sigma \underline{a})\,,$$

and by Cramér - Wold, this yields

$$\sqrt{n}(\widehat{\underline{p}}_n - \underline{p}_n) \to_d N_k(0, \Sigma)\,.$$

2. Suppose that $\underline{N}_n \sim$ $\text{Mult}_k(n, \underline{p})$ and $\widehat{p} = \underline{N}_n/n$. Define a family of functions ϕ_s for $-1 \le s \le 2$ by

$$\phi_s(x) = \frac{1 - s + sx - x^s}{s(1-s)}, \qquad x \in \mathbb{R}^+, \quad s \neq 0, 1,$$

and define $\phi_1(x) = x(\log x - 1) + 1$, $\phi_0(x) = \log(1/x) + x - 1$. Now define a family of statistics for testing $H:\ \underline{p} = \underline{p}_0$ versus $K:\ \underline{p} \neq \underline{p}_0$ by

$$T_n(s) \equiv 2n \sum_{j=1}^k p_{0j} \phi_s\left(\frac{\hat{p}_j}{p_{0j}}\right).$$

(a) Show that $T_n(s)$ reduces to the following statistics discussed in class:
(i) $T_n(2)$ is the Pearson chi-square statistic Q_n; (ii) $T_n(1)$ is $2\log\lambda_n$ where λ_n is the likelihood ratio statistic; (iii) $T_n(-1)$ is Neyman's version of the chi-square statistic, Q_n^{Neyman}; and (iv) $T_n(1/2)$ is the Hellinger statistic H_n^2.
(b) Show that $n^{-1}T_n(s)$ converges in probability to a deterministic limit $t(s)$ under a general $\underline{p}$, and identify the limit explicitly in cases (i) - (iv) of part (a). Do any of these limiting parameters have names?
(c) Find the limiting distribution of $T_n(1/2)$ under the null hypothesis H.
Notes: This problem is related to the statistics treated in Cressie and Read, JRSS

B 46 (1984), 440 - 464, and also to the "transformed" chi-square statistics discussed in Ferguson, ACILST, pages 59 and 66. See also Jager and Wellner, Ann. Statist. 35 (2007), 2018-2053 for related material in a different vein.

Solution: (a) (i) When $s = 2$, $\phi_2(x) = (-1 + 2x - x^2)/(-2) = 2^{-1}(x-1)^2$, so

$$\begin{aligned} T_n(2) &= 2n \sum_{j=1}^{k} p_{0j}\phi_2(\hat{p}_j/p_{j0}) = n \sum_{j=1}^{k} p_{0j}(\hat{p}_j/p_{0j} - 1)^2 \\ &= n \sum_{j=1}^{k} \frac{(\hat{p}_j - p_{0j})^2}{p_{0j}} = Q_n. \end{aligned}$$

(ii) When $s = 1$, $\phi_1(x) = x(\log x - 1) + 1$, so

$$\begin{aligned} T_n(1) &= 2n \sum_{j=1}^{k} p_{0j}\phi_2(\hat{p}_j/p_{j0}) = 2n \sum_{j=1}^{k} \{\hat{p}_j(\log(\hat{p}_j/p_{0j}) - 1) + p_{0j}\} \\ &= 2\sum_{j=1}^{k} N_j \log(N_j/(np_{0,j})) = 2\log\lambda_n, \end{aligned}$$

where $\lambda_n = \prod_{j=1}^{k}(N_j/(np_{0,j}))^{N_j}$ is the likelihood ratio statistic.
(iii) When $s = -1$, $\phi_{-1}(x) = (2 - x - x^{-1})/(-1 \cdot 2) = 2^{-1}(x-1)^2/x$, so

$$\begin{aligned} T_n(-1) &= 2n \sum_{j=1}^{k} p_{0j}\phi_{-1}(\hat{p}_j/p_{j0}) = n \sum_{j=1}^{k} (\hat{p}_j - p_{0j})^2/\hat{p}_j \\ &= \sum_{j=1}^{k} (N_j - np_{0j})^2/N_j = Q_n^{Neyman}. \end{aligned}$$

(iv) When $s = 1/2$, $\phi_{1/2}(x) = 2^{-1} + 2^{-1}x - x^{1/2}$, so

$$\begin{aligned} T_n(1/2) &= 2n \sum_{j=1}^{k} p_{0j}\phi_{1/2}(\hat{p}_j/p_{j0}) = 2n \sum_{j=1}^{k} \{2^{-1}p_{0,j} + 2^{-1}\hat{p}_j - \sqrt{\hat{p}_j p_{0j}}\} \\ &= 2n(1 - \sum_{j=1}^{k} \sqrt{\hat{p}_j p_{0j}}) = 2nH^2(\underline{\hat{p}}, \underline{p}_0) \end{aligned}$$

where $H^2(\underline{\hat{p}}, \underline{p}_0) = 2^{-1}\sum_{j=1}^{k}(\sqrt{\hat{p}_j} - \sqrt{p_{0j}})^2$.
(b) In general we have, by the continuous mapping theorem since $g_s(x) = \sum_{j=1}^{k} p_{0j}\phi_s(x_j/p_{0j})$ is continuous in $x = (x_1, \ldots, x_k) \in (0,1]^k$ for each $s \in [-1, 2]$ and $\underline{\hat{p}} \to_p \underline{p}$,

$$n^{-1}T_n(s) \to_p 2g_s(\underline{p}) = 2\sum_{j=1}^{k} p_{0j}\phi_s(p_j/p_{0j}) \equiv t(s) = t(s; \underline{p}, \underline{p}_0).$$

Much as in (a), it follows that

$$t(s;\underline{p},\underline{p}_0) \; = \; \begin{cases} \sum_{j=1}^k (p_j - p_{j0})^2/p_{j0} \equiv q(\underline{p},\underline{p}_0), & \text{when } s = 2, \\ 2\sum_{j=1}^k p_j \log(p_j/p_{0j}) \equiv 2K(\underline{p},\underline{p}_0), & \text{when } s = 1, \\ 2\sum_{j=1}^k p_{0j} \log(p_{0j}/p_j) \equiv 2K(\underline{p}_0,\underline{p}), & \text{when } s = 0, \\ \sum_{j=1}^k (p_j - p_{0j})^2/p_j \equiv q(\underline{p}_0,\underline{p}), & \text{when } s = -1, \\ 2(1-\sum_{j=1}^k \sqrt{p_j p_{j0}}) \equiv 2H^2(\underline{p},\underline{p}_0), & \text{when } s = 1/2. \end{cases}$$

Here $q(\underline{p},\underline{p}_0)$ is the "chi-square discrepancy" between $\underline{p}$ and $\underline{p}_0$, $K(\underline{p}_0,\underline{p})$ is the "Kullback-Leibler discrepancy" between $\underline{p}_0$ and $\underline{p}$, and $H^2(\underline{p},\underline{p}_0)$, is the Hellinger distance (metric!) between $\underline{p}$ and $\underline{p}_0$.

(c) We know from the multivariate CLT that when $H: p = p_0$ is true, then

$$\sqrt{n}(\underline{\hat{p}} - \underline{p}_0) \to_d N_k(0, A)$$

where $A = \text{diag}(\underline{p}_0) - \underline{p}_0\underline{p}_0^T$. Thus with $g(\underline{u}) \equiv (\sqrt{u_1}, \ldots, \sqrt{u_k})^T$, and hence

$$\nabla g(\underline{p}_0) = 2^{-1}\text{diag}(1/\sqrt{\underline{p}_0}),$$

it follows by the delta method that

$$2\sqrt{n}(\sqrt{\underline{\hat{p}}} - \sqrt{\underline{p}_0}) \to_d \underline{Z} \sim N_k(0, I - \sqrt{\underline{p}_0}\sqrt{\underline{p}_0}^T).$$

Therefore, by the continuous mapping theorem,

$$4n\sum_{j=1}^k (\sqrt{\hat{p}_j} - \sqrt{p_{0j}})^2 \to_d \underline{Z}^T\underline{Z} = |\underline{Z}|^2.$$

As we have seen in class, the distribution of the random variable $|\underline{Z}|^2$ is χ^2_{k-1}.

Notes: In fact, $T_n(s) \to_d \chi^2_{k-1}$ under $H: \underline{p} = \underline{p}_0$ for all $s \in [-1, 2]$. Furthermore, under $\underline{p}_n = \underline{p}_0 + \underline{c}n^{-1/2}$, it can be shown that $T_n(s) \to_d \chi^2_{k-1}(\delta)$ with $\delta = \sum_{j=1}^k c_j^2/p_{0j}$ for all $s \in [-1, 2]$. But as seen in (b), the statistics $n^{-1}T_n(s)$ converge to different natural parameters when $\underline{p} \neq \underline{p}_0$ holds, and hence yield different tests.

3. Ferguson, ACILST, problem 4, page 55, modified slightly: suppose that the sample sizes (of X's and Y's) are m and n respectively, and that $m/(m+n) \to \lambda \in (0,1)$.

Solution: Since $S_X^2 \to_p \sigma^2$ and $S_Y^2 \to_p \tau^2$ as $m \wedge n \to \infty$, $S_Y^2/S_X^2 \to_p \tau^2/\sigma^2$ as $m \wedge n \to \infty$. Furthermore, with $N \equiv m + n$ and using the assumption that μ_4 , $\nu_4 < \infty$,

$$\begin{aligned} \sqrt{\frac{mn}{2N}}\left(\frac{S_Y^2}{S_X^2} - \frac{\tau^2}{\sigma^2}\right) &= \sqrt{\frac{m}{N}}\sqrt{\frac{n}{2}}\frac{(S_Y^2 - \tau^2)}{S_X^2} + \sqrt{\frac{n}{N}}\sqrt{\frac{m}{2}}\left(\frac{1}{S_X^2} - \frac{1}{\sigma^2}\right)\tau^2 \\ &= \sqrt{\frac{m}{N}}\sqrt{\frac{n}{2}}\left(\frac{S_Y^2}{\tau^2} - 1\right)\frac{\tau^2}{S_X^2} - \sqrt{\frac{n}{N}}\sqrt{\frac{m}{2}}\left(\frac{S_X^2}{\sigma^2} - 1\right)\frac{\tau^2}{S_X^2} \\ &\to_d \sqrt{\lambda}\frac{\tau^2}{\sigma^2}N(0, 1 + \gamma_{Y2}/2) - \sqrt{1-\lambda}\frac{\tau^2}{\sigma^2}N(0, 1 + \gamma_{X2}/2) \\ &\qquad \text{where the two normal rv's are independent} \\ &\sim \frac{\tau^2}{\sigma^2}N(0, 1 + (1/2)(\lambda\gamma_{Y2} + \overline{\lambda}\gamma_{X2})) \\ &\equiv \frac{\tau^2}{\sigma^2}N(0, 1 + (1/2)\overline{\gamma}_2) \end{aligned}$$

where $\overline{\gamma}_2 \equiv \lambda\gamma_{Y2} + \overline{\lambda}\gamma_{X2}$ if $\lambda_N \to \lambda$ as $m \wedge n \to \infty$ since the two normal rv's are independent.

When $\sigma^2 = \tau^2$ and the X's and Y's are normal, so $\gamma_{X2} = \gamma_{Y2} = 0$, it follows that $S_Y^2/S_X^2 \equiv F$ has an $F_{n-1,m-1}$ distribution, so that

$$\begin{aligned}
\alpha &= P_{\sigma^2=\tau^2}(F \geq F_{n-1,m-1,\alpha}) \\
&= P_{\sigma^2=\tau^2}\left(\sqrt{\frac{mn}{2N}}\left(\frac{S_Y^2}{S_X^2} - 1\right) \geq \sqrt{\frac{mn}{2N}}(F_{n-1,m-1,\alpha} - 1)\right) \\
&\to P(N(0,1) \geq z_\alpha) = \alpha;
\end{aligned}$$

that is, we must have

$$\sqrt{\frac{mn}{2N}}(F_{n-1,m-1,\alpha} - 1)) \to z_\alpha.$$

Thus, when the X's and Y's are *not normal*,

$$\begin{aligned}
P_{\sigma^2=\tau^2}(F \geq F_{n-1,m-1,\alpha}) &= P_{\sigma^2=\tau^2}\left(\sqrt{\frac{mn}{2N}}(\frac{S_Y^2}{S_X^2} - 1) \geq \sqrt{\frac{mn}{2N}}(F_{n-1,m-1,\alpha} - 1)\right) \\
&\to P(N(0, 1 + \frac{\overline{\gamma}_2}{2}) \geq z_\alpha) \\
&= P(N(0,1) \geq z_\alpha/\sqrt{1 + \overline{\gamma}_2/2}) \\
&= 1 - \Phi(\frac{z_\alpha}{\sqrt{1+\overline{\gamma}_2/2}}) \begin{cases} > \alpha & \text{if } \overline{\gamma}_2 > 0 \\ < \alpha & \text{if } \overline{\gamma}_2 < 0 \end{cases} .
\end{aligned}$$

Recall from problem 2.5 that $\gamma_2 \in [-2, \infty)$ always; and note that

$$\lim_{\overline{\gamma}_2 \to -2} \{1 - \Phi(\frac{z_\alpha}{\sqrt{1+\overline{\gamma}_2/2}})\} = 0$$

while

$$\lim_{\overline{\gamma}_2 \to \infty} \{1 - \Phi(\frac{z_\alpha}{\sqrt{1+\overline{\gamma}_2/2}}) = 1/2\,.$$

4. Suppose that $\underline{N}_n = (N_{11}, N_{12}, N_{21}, N_{22}) \sim \text{Mult}_4(n, \underline{p})$ where $\underline{p} = (p_{11}, p_{12}, p_{21}, p_{22})$ where $\sum_{i=1}^2 \sum_{j=1}^2 p_{ij} = 1$. (Thus $\underline{N}_n$ is the sum of n independent $\text{Mult}_4(1, \underline{p})$ random vectors $\{\underline{Y}_i\}_{i=1}^n$.) Since there are really just three independently varying parameters for this problem, it is often useful to re-express the cell probabilities in terms of two marginal probabilities, say $p_{1\cdot} = p_{11} + p_{12}$ and $p_{\cdot 1} = p_{11} + p_{21}$, and ψ, the log of the odds-ratio, defined by

(2) $$\psi \equiv \log \frac{p_{21}/p_{22}}{p_{11}/p_{12}} = \log \frac{p_{12}p_{21}}{p_{11}p_{22}}\,.$$

You may use the fact that $\psi = 0$ if and only if independence holds for the 2×2 table (i.e. $p_{ij} = p_{i\cdot}p_{\cdot j}$ for $i, j = 1, 2$).
(a) Suggest an estimator of ψ, say $\widehat{\psi}$.
(b) Show that the estimator you proposed in (a) is asymptotically normal and compute the asymptotic variance of your estimator.

Solution: (a) An obvious estimator of ψ is

$$\widehat{\psi} = \log \frac{\widehat{p}_{12}\widehat{p}_{21}}{\widehat{p}_{11}\widehat{p}_{22}}$$

where $\underline{\widehat{p}} = \underline{N}/n$.

(b) Now $\widehat{\psi} = g(\underline{\widehat{p}})$ where $g(\underline{p})$ is given in (2) and is differentiable with derivative

$$\nabla g(\underline{p}) = (-1/p_{11}, 1/p_{12}, 1/p_{21}, -1/p_{22})$$

and, by the multivariate CLT,

$$\sqrt{n}(\underline{\widehat{p}} - \underline{p}) \to_d Z \sim N_4(0, \Sigma)$$

where $\Sigma = \text{diag}(\underline{p}) - \underline{p}\underline{p}^T$. Thus the delta method (or g'-theorem) yields

$$\begin{aligned} \sqrt{n}(\widehat{\psi} - \psi) &= \sqrt{n}(g(\underline{\widehat{p}}) - g(\underline{p})) \\ &\to_d \nabla g(\underline{p}) Z \sim N(0, \nabla g^T \Sigma \nabla g) = N(0, V^2(\underline{p})) \end{aligned}$$

where

$$V^2(\underline{p}) = \frac{1}{p_{11}} + \frac{1}{p_{12}} + \frac{1}{p_{21}} + \frac{1}{p_{22}}.$$

5. This is a continuation of problem 3. One standard test of independence in the 2×2 table is the test based on a Pearson-type chi-square statistic.

(a) Write down the chi-square statistic Q_n for this problem, state its asymptotic distribution under the null hypothesis, and explain briefly why the claimed result holds.

(b) Suppose that the alternative hypothesis holds. Show that under the alternative hypothesis $n^{-1} Q_n \to_p$ some constant q and compute q as explicitly as possible.

(c) Find the asymptotic distribution of Q_n under local alternatives of the form $\psi_n = tn^{-1/2}$; i.e. $\underline{p}_n \equiv (p_{11,n}, p_{12,n}, p_{21,n}, p_{22,n}) = \underline{p}_0 + \underline{c} n^{-1/2}$ where

$$\psi_0 \equiv \log\left(\frac{p_{21,0} p_{12,0}}{p_{11,0} p_{22,0}}\right) = 0$$

and $\underline{1}'\underline{c} = 0$.

(d) Suppose that $n = 30$, $\alpha = .02$, and the true $\underline{p}$ is $\underline{p} = (.3, .2, .1, .4)$. Give an approximation to the power of the chi-square test at this particular alternative.

Solution: (a) The chi-square statistic for testing independence in a 2×2 table is

$$\begin{aligned} Q_n &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(N_{ij} - n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}} \\ &= \frac{(N_{11}N_{22} - N_{12}N_{21})^2}{n^3} \sum_{i,j} \left\{ \frac{1}{\hat{p}_{i\cdot}\hat{p}_{\cdot j}} \right\} \\ &= \frac{(N_{12}N_{21} - N_{11}N_{22})^2}{n^3} \frac{1}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\ &= \frac{n\{\exp(\hat{\psi}_n) - 1\}^2 (\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\ &= \frac{\{\sqrt{n}[\exp(\hat{\psi}_n) - 1]\}^2 (\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \end{aligned}$$

$$
\begin{aligned}
&\to_d \quad [N(0,V^2)]^2 \frac{[p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1})]^2}{p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1})} \\
&= \quad [N(0,V^2)]^2 p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1}) = [N(0,1)]^2 \stackrel{d}{=} \chi_1^2
\end{aligned}
$$

by the delta method or g' theorem and result of problem 3 where we have repeatedly used the fact that $p_{ij} = p_{i\cdot}p_{\cdot j}$ under the null hypothesis of independence.
(b) When the alternative hypothesis holds, then the above argument shows that

$$
\begin{aligned}
n^{-1}Q_n &= \frac{(N_{12}N_{21} - N_{11}N_{22})^2}{n^4} \frac{1}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\
&= \frac{(\hat{p}_{12}\hat{p}_{21} - \hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\
&\to_p \frac{(p_{12}p_{21} - p_{11}p_{22})^2}{p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1})}
\end{aligned}
$$

where $p_{1\cdot} = p_{11} + p_{12}$ and $p_{\cdot 1} = p_{11} + p_{21}$.
(c) Under local alternatives with $\psi_n = tn^{-1/2}$ for $t \neq 0$, the argument in (a) repeated (but using the Liapunov CLT) yields

$$
\begin{aligned}
\sqrt{n}(\widehat{\psi}_n - 0) &= \sqrt{n}(\widehat{\psi}_n - \psi_n) + \sqrt{n}(\psi_n - 0) \\
&= \sqrt{n}(g(\underline{\widehat{p}}) - g(\underline{p}_n)) + t \\
&\to_d \nabla g(\underline{p}_0)Z + t \sim N(t, \nabla g^T \Sigma \nabla g) = N(t, V^2(\underline{p}_0))
\end{aligned}
$$

where

$$
V^2(\underline{p}_0) = \frac{1}{p_{11,0}} + \frac{1}{p_{12,0}} + \frac{1}{p_{21,0}} + \frac{1}{p_{22,0}} = \frac{1}{p_{1\cdot,0}(1-p_{1\cdot,0})p_{\cdot 1,0}(1-p_{\cdot 1,0})},
$$

and hence, by the delta-method again,

$$
\sqrt{n}(\exp(\widehat{\psi}_n) - 1) \quad \to_d \quad \nabla g(\underline{p}_0)Z + t \sim N(t, \nabla g^T \Sigma \nabla g) = N(t, V^2(\underline{p}_0)).
$$

This implies, via the same development as in (a), that under $\underline{p}_n$ we have

$$
\begin{aligned}
Q_n &= \frac{n\{\exp(\hat{\psi}_n) - 1\}^2 (\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\
&= \frac{\{\sqrt{n}[\exp(\hat{\psi}_n) - 1]\}^2 (\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\
&\to_d [N(t, V^2(\underline{p}_0))]^2 p_{1\cdot,0}(1-p_{1\cdot,0})p_{\cdot 1,0}(1-p_{\cdot 1,0}) \\
&= [N(t\sqrt{c}, 1)]^2 \stackrel{d}{=} \chi_1^2(\delta)
\end{aligned}
$$

where $\delta = ct^2$ and $c \equiv p_{1\cdot,0}(1-p_{1\cdot,0})p_{\cdot 1,0}(1-p_{\cdot 1,0})$.
(d) When $n = 30$, $\alpha = .02$, and the true $\underline{p}$ is $\underline{p} = (.3, .2, .1, .4)$, we calculate $p_{1\cdot} = 1 - p_{\cdot 1} = .5$, $p_{\cdot 1} = .4$ (so that $c = p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1}) = (.5)^2(.4)(.6) = .06$),

$$
t_n \equiv \sqrt{n} \log \frac{p_{12}p_{21}}{p_{11}p_{22}} = -9.814... \, .
$$

Thus $\delta = (9.814)^2(.06) = 5.779...$, and an approximation to the power of our test is given by

$$
P(\chi_1^2(5.779...) > \chi_{1,.02}^2) = P(\chi_1^2(5.779...) > 5.412...) = .531... \, .
$$