Statistics 581, Problem Set 4 Solutions

Wellner; 10/27/2010

1. Suppose that $X_1, X_2, \ldots$ are i.i.d. (μ, σ^2) with $\mu_4 < \infty$. Let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ be the sample mean and sample variance respectively.
(a) Show that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}.$$

(b) Suppose $\mu \neq 0$. Use (a) to find the limiting distribution of the sample *coefficient of variation* $C_n \equiv S_n/\overline{X}_n$; i.e. show that $\sqrt{n}(C_n - c) \to_d N(0, V^2)$ with $c \equiv \sigma/\mu$ and find V^2.

Solution: (a) Since $S_n^2 = n^{-1}\sum_{i=1}^n (X_i - \mu)^2 + o_p(1/\sqrt{n})$, we have

$$\begin{aligned}
\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \begin{pmatrix} X_i - \mu \\ (X_i - \mu)^2 - \sigma^2 \end{pmatrix} + o_p(1) \\
&\to_d \underline{Z} \sim N_2(0, \Sigma)
\end{aligned}$$

by the multivariate CLT where Σ is as given above.
(b) The function $g(u, v) = \sqrt{v}/u$ is differentiable at points (u, v) with $u \neq 0$, and the derivative is $\nabla g(u, v) = (-u^{-2}\sqrt{v}, (1/2)v^{-1/2}u^{-1})$ so that $\nabla g(\mu, \sigma^2) = (-\mu^{-2}\sigma, (1/2)\sigma^{-1}\mu^{-1}) = (\sigma/\mu)(-1/\mu, (1/2\sigma^2))$. Hence it follows from the delta method (g' theorem) that

$$\begin{aligned}
\sqrt{n}(C_n - c) &= \sqrt{n}(g(\overline{X}_n, S_n^2) - g(\mu, \sigma^2)) \\
&\to_d \nabla g \cdot \underline{Z} \sim N(0, \nabla g^T \Sigma \nabla g)
\end{aligned}$$

and it is easy to calculate that

$$\begin{aligned}
\nabla g^T \Sigma \nabla g &= c^2 \left\{ \frac{\sigma^2}{\mu^2} - \frac{\mu_3}{2\mu\sigma^2} + \frac{1}{2}(1 + \frac{\gamma_2}{2}) \right\} \\
&= c^2(c^2 - c\gamma_1 + \frac{1}{2}(1 + \frac{\gamma_2}{2}))
\end{aligned}$$

where $\gamma_1 \equiv \mu_3/\sigma^3$. Note that when the X_i's are normal (so $\gamma_1 = \gamma_2 = 0$), this reduces to $c^2(c^2 + 1/2)$.

2. (a) Suppose that $\underline{N}_n \sim \text{Mult}_k(n, \underline{p})$ 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 $\underline{Z}_n \to_d \underline{Z}$ 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 $Q_n \to_d \chi^2_{k-1}(\delta)$ with $\delta = \sum_{j=1}^k c_j^2/p_{0,j}$ under the local alternative $\underline{p}_n$.)

(b) (Ferguson, *A Course in Large Sample Theory*, page 65.) In a multinomial experiment with sample size $n = 100$ and 3 cells with null hypothesis $H_0 : \underline{p}_0 = (.25, .5, .25)$, what is the approximate power at the alternative $\underline{p} = (.2, .6, .2)$ when the level of significance is $\alpha = .05$? $\alpha = .01$? How large a sample size is need to achieve power 0.9 at this alternative when $\alpha = .05$? $\alpha = .01$?

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

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

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

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

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

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

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

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

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

$$\sigma_n^2 = \sum_1^n \sigma_{ni}^2 = \underline{a}^T(\text{diag}(\underline{p}_n) - \underline{p}_n\underline{p}_n^T)\underline{a} \to \underline{a}^T\Sigma\underline{a}$$

while

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

where

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

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

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

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 implies

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

(b) Now

$$n^{1/2}(\underline{p} - \underline{p}_0) = 10((.2, .6, .2) - (.25, .5, .25)) = 10(-.05, .10, -0.05) = (-.5, 1, -.5)\,,$$

so the non-centrality parameter is

$$\delta = \frac{(.5)^2}{.25} + \frac{1^2}{.5} + \frac{(.5)^2}{.25} = 1 + 2 + 1 = 4\,.$$

Thus the approximate power via $\chi_2^2(\delta)$ is

$$P(\chi_2^2(4) \ge \chi_{2,.05}) = P(\chi_2^2(4) \ge 5.991) = 0.415427\,, \qquad \text{when} \quad \alpha = .05\,,$$

and

$$P(\chi_2^2(4) \ge \chi_{2,.01}) = P(\chi_2^2(4) \ge 9.210) = 0.203948 \qquad \text{when} \quad \alpha = .01\,,$$

Now we want to find n so that

$$P(\chi_2^2(\delta_n) \ge 5.991) = .90$$

where

$$\delta_n = n \left\{ \frac{(.05)^2}{.25} + \frac{(.1)^2}{.5} + \frac{(.05)^2}{.25} \right) = n/25.$$

In this case we find that $\delta_n = n/25 = 12.6539$, so that $n = 25 \cdot 12.6539 \approx 316$. When $\alpha = .01$ we find that $\delta_n = n/25 = 17.4267$ so that $n = 25 \cdot 17.4267/(.04) \approx 436$.

The alternative approximation to power that we derived in class is

$$\begin{aligned} P_p(Q_n \ge \chi^2_{k-1,\alpha}) &= P_p\left(\sqrt{n}(n^{-1}Q_n - q) \ge \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &\doteq P\left(N(0, d^T A d) \ge \sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)\right) \\ &= 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{k-1,\alpha} - q)/\sqrt{d^T A d}) \end{aligned}$$

where $d \equiv 2\text{diag}(1/p_0)(p - p_0)$, $A = \text{diag}(p) - pp^T$, and $q = \sum_{j=1}^k (p_j - p_{j0})^2/p_{j0}$. In the present case I calculate $q = 1/25$, $d = .4(-1, 1, -1)^T = (2/5)(-1, 1, -1)$, and

$$A = \text{diag}(p) - pp^T = \frac{1}{25}\begin{pmatrix} 4 & -3 & -1 \\ -3 & 6 & -3 \\ -1 & -3 & 4 \end{pmatrix}$$

so that $d^T A d = 4 \cdot 24/5^4$. Thus the approximation becomes

$$P_p(Q_n \geq \chi^2_{2,\alpha}) \doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,\alpha} - 1/25)/(2\sqrt{24}/5^2)).$$

When I calculate I get

$$\begin{aligned} P_p(Q_n \geq \chi^2_{2,.05}) &\doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,.05} - 1/25)/(2\sqrt{24}/5^2)) = 0.30568 \\ P_p(Q_n \geq \chi^2_{2,.05}) &\doteq 1 - \Phi(\sqrt{n}(n^{-1}\chi^2_{2,.01} - 1/25)/(2\sqrt{24}/5^2)) = 0.0918505, \end{aligned}$$

which are both somewhat lower than suggested by the non-central chi-square approximation. A Monte-Carlo study not shown here shows that the non-central chi-square approximation is quite accurate in this case. I suspect that the fixed alternative limit theorem and resulting normal approximation to power will do better for more extreme alternatives with a larger number of cells, but I have not carried out a thorough study.

3. Suppose the same set-up as in the chi-square testing situation considered in lecture in class but now, for testing $H_0 : \underline{p} = \underline{p}_0$ versus $K_0 : \underline{p} \neq \underline{p}_0$, instead of the chi-square statistic Q_n, consider the test statistic given by

$$H_n^2 \equiv 4n \sum_{i=1}^{k} (\sqrt{\hat{p}_i} - \sqrt{p_{i0}})^2.$$

The statistic H_n^2 is $4n$ times the square of the *Hellinger distance* between $\underline{\hat{p}}$ and $\underline{p}_0$.
(a) Find the limiting distribution of H_n^2 under the null hypothesis H_0.
(b) Find the limit of $n^{-1}H_n^2$ under fixed alternatives $\underline{p} \neq \underline{p}_0$ in K_0, and use this to show that the test based on H_n^2 is consistent against K_0.
(c) Find the limiting distribution of H_n^2 under local alternatives $\underline{p}_n = \underline{p}_0 + \underline{c}/\sqrt{n}$, and use this to approximate the power of this test. Compare the (local asymptotic) power of this test to the chi-square test.

Solution: (a) Let $\underline{Z}_n \equiv \sqrt{n}(\underline{\hat{p}}_n - \underline{p}_0)$ Then $\underline{Z}_n \to_d \underline{Z} \sim N_k(0, \Sigma)$ with $\Sigma = \text{diag}(\underline{p}_0) - \underline{p}_0 \underline{p}_0^T$. Thus, by the delta - method,

$$\begin{aligned} \underline{Y}_n &\equiv 2\sqrt{n}(\sqrt{\underline{\hat{p}}_n} - \sqrt{\underline{p}_0}) \\ &\to_d \text{diag}(1/\sqrt{\underline{p}_0})\underline{Z} \equiv \underline{Y} \sim N_k(0, I - \sqrt{\underline{p}_0}\sqrt{\underline{p}_0^T}) \end{aligned}$$

Hence, by the continuous mapping theorem,

$$H_n^2 = \underline{Y}_n^T \underline{Y}_n \to_d \underline{Y}^T \underline{Y}.$$

It remains to answer the question: what is the distribution of $\underline{Y}^T \underline{Y}$? This goes just exactly as in the case of the limit for the chi-square statistic Q_n. Let Γ be an orthogonal matrix with first row $\sqrt{\underline{p}_0}$. Then

$$\Gamma \underline{Y} \sim N_k(0, \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}),$$

which has first coordinate 0, and the remaining $k-1$ coordinates are iid $N(0,1)$. Further, $\Gamma^T\Gamma = I$ and hence

$$\underline{Y}^T\underline{Y} = \underline{Y}^T\Gamma^T\Gamma\underline{Y} = (\Gamma\underline{Y})^T(\Gamma\underline{Y}) \sim \chi^2_{k-1}.$$

Thus $H_n^2 \rightarrow_d \underline{Y}^T\underline{Y} \sim \chi^2_{k-1}$.
(b) Under fixed $\underline{p} \neq \underline{p}_0$, $\hat{p}_n \rightarrow_{a.s.} p$. Hence by the continuous mapping theorem

$$\begin{aligned} n^{-1}H_n^2 &= 4\sum_{j=1}^k \left\{\sqrt{\hat{p}_j} - \sqrt{p_{j0}}\right\}^2 \\ &\rightarrow_{a.s.} 4\sum_{j=1}^k (\sqrt{p_j} - \sqrt{p_{j0}})^2 \\ &= 4d_H^2(p, p_0) > 0. \end{aligned}$$

Therefore, under $p \neq p_0$, $H_n^2 \rightarrow_{a.s.} \infty$, and hence

$$P_p(H_n^2 \geq \chi^2_{k-1,\alpha}) \rightarrow 1.$$

(c) Under local alternatives, Liapunov's CLT, the Cramér - Wold device, and the delta method, yield

$$\begin{aligned} \underline{Y}_n &= 2\sqrt{n}(\sqrt{\underline{\hat{p}}_n} - \sqrt{\underline{p}_n}) + 2\sqrt{n}(\sqrt{\underline{p}_n} - \sqrt{\underline{p}_0}) \\ &\rightarrow_d \underline{Y} + \text{diag}(1/\sqrt{\underline{p}})\underline{c} \\ &\equiv \underline{Y} + \underline{\mu} \\ &\sim N_k(\underline{\mu}, I - \sqrt{p_0}\sqrt{p_0}^T). \end{aligned}$$

Now with Γ as in part (a)

$$\Gamma(\underline{Y} + \underline{\mu}) = \Gamma\underline{Y} + \Gamma\underline{\mu}) = \Gamma\underline{Y} + \underline{b}$$

where the first coordinate of $\underline{b}$ is 0. Thus $\Gamma\underline{Y} + \underline{b}$ has first coordinate 0, and the remaining $k-1$ coordinates are independent $N(b_i, 1)$. Hence

$$\begin{aligned} (\underline{Y} + \underline{\mu})^T(\underline{Y} + \underline{\mu}) &= (\Gamma\underline{Y} + \underline{b})^T(\Gamma\underline{Y} + \underline{b}) \\ &\sim \chi^2_{k-1}(\underline{b}^T\underline{b}) = \chi^2_{k-1}(\sum_{j=1}^k c_i^2/p_{i0}) \end{aligned}$$

Thus the local asymptotic power of the test based on the Hellinger statistics H_n^2 is the same as that of the chi-square statistic Q_n.

4. Suppose that $Y_i = \alpha + \theta'(x_i - \overline{x}) + \epsilon_i$, $i = 1, \ldots, n$, where $\epsilon_i \sim (0, \sigma^2)$ are i.i.d. and the x_i's are known vectors in R^k. Equivalently, $\underline{Y} = X\underline{\beta} + \underline{\epsilon}$ where

$$X^T = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 - \overline{x} & x_2 - \overline{x} & \cdots & x_n - \overline{x} \end{pmatrix}$$

so that X is an $n \times (k+1)$ matrix. Let $\underline{\hat{\beta}}$ be the least squares estimator of $\underline{\beta} = (\alpha, \theta')'$; i.e. $\underline{\hat{\beta}} = (X^T X)^{-1} X^T \underline{Y}$. Suppose that $n^{-1}(X^T X) \to D$ where D is positive definite.

(a) What additional condition(s) do you need to impose to prove that

$$\sqrt{n}(\hat{\beta}_n - \beta) \to_d N_{k+1}(0, \text{“}\textit{something}\text{”})\,?$$

(b) Find “something” in part (a).

Solution: (a) Let $\underline{a} \in R^{k+1}$. Now

$$\begin{aligned} \hat{\beta} &= (X^T X)^{-1} X^T Y \\ &= (X^T X)^{-1} X^T (X\beta + \epsilon) \\ &= \beta + (X^T X)^{-1} X^T \epsilon\,, \end{aligned}$$

so

$$\sqrt{n}(\hat{\beta} - \beta) = \sqrt{n}(X^T X)^{-1} X^T \epsilon \equiv B_n \epsilon$$

where $B_n \equiv \sqrt{n}(X^T X)^{-1} X^T$ is a $(k+1) \times n$ matrix. Thus

$$\begin{aligned} a^T(\sqrt{n}(\hat{\beta} - \beta)) &= a^T B_n \epsilon \equiv b_n^T \epsilon \\ &= \sum_{i=1}^n b_{ni}\epsilon_i \equiv \sum_{i=1}^n X_{ni} \end{aligned}$$

where $b_n^T \equiv a^T B_n$ is an $1 \times n$ vector. Now we compute

$$\mu_{ni} = E(X_{ni}) = b_{ni} \cdot 0, \qquad \sigma_{ni}^2 = Var(X_{ni}) = b_{ni}^2 \sigma^2\,,$$

and hence, using the hypothesized convergence of $n^{-1} X^T X \to D$ in the last line,

$$\begin{aligned} \sigma_n^2 &= \sigma^2 \sum_{i=1}^n b_{ni}^2 = \sigma^2 b_n^T b_n \\ &= \sigma^2 a^T B_n B_n^T a = n\sigma^2 a^T (X^T X)^{-1}(X^T X)(X^T X)^{-1} a \\ &= \sigma^2 a^T (n^{-1} X^T X)^{-1} a \to \sigma^2 a^T D^{-1} a \equiv V^2(a) > 0 \end{aligned}$$

since D is nonsingular. To establish asymptotic normality of $a^T(\sqrt{n}(\hat{\beta} - \beta))/\sigma_n$, it remains to verify the Lindeberg condition: namely

(2) $$\frac{1}{\sigma_n^2} \sum_{i=1}^n E\left\{ |X_{ni}|^2 1_{[|X_{ni}| > \delta\sigma_n]} \right\} \to 0$$

for every $\delta > 0$. But, as we have seen before, this holds if

(3) $$\max_{1 \le i \le n} |b_{ni}| \to 0 \qquad \text{as} \qquad n \to \infty\ :$$

the left side of (2) is bounded as follows:

$$\begin{aligned} &\frac{1}{\sigma_n^2} \sum_{i=1}^n b_{ni}^2 E\left\{ \epsilon_1^2 1_{[|\epsilon_1| > \delta\sigma_n/|b_{ni}|]} \right\} \\ &\le \frac{1}{\sigma^2} E\left\{ \epsilon_1^2 1_{[|\epsilon_1| > \delta\sigma_n/\max_{1\le i\le n}|b_{ni}|]} \right\} \\ &\to \frac{1}{\sigma^2} \cdot 0 = 0 \end{aligned}$$

by (3), $E(\epsilon_1^2) < \infty$, and the dominated convergence theorem. Thus it follows from the Lindeberg-Feller CLT that

$$a^T(\sqrt{n}(\hat\beta - \beta))/\sigma_n \to_d N(0,1),$$

and since $\sigma_n^2 \to \sigma^2 a^T D^{-1} a$, this implies that

$$a^T(\sqrt{n}(\hat\beta - \beta)) \to_d N(0, a^T(\sigma^2 D^{-1})a)\,,$$

which in turn, via the Cramér-Wold device, implies

$$\sqrt{n}(\hat\beta - \beta) \to_d N_{k+1}(0, \sigma^2 D^{-1})\,.$$

5. 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

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

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 (4) 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

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

6. **Optional bonus problem 1:** This is a continuation of problem 5 above. 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.

7. **Optional bonus problem 2:** Suppose that $(X_i - \mu)/\sigma$, $i = 1, \ldots, m$ and $(Y_j - \nu)/\tau$, $j = 1, \ldots, n$ are iid $(0, 1, \mu_4 < \infty)$ (thus γ_2 is the same for the two populations), and let S_X^2 and S_Y^2 denote the sample variances of the X's and Y's respectively. The classical $F-$test based on the assumption that all the standardized X's and Y's are $N(0,1)$ rejects $H_0 : \tau \le \sigma$ in favor of $H_1 : \tau > \sigma$ if $F \equiv S_Y^2/S_X^2 > F_{n-1,m-1,\alpha}$. Assuming that $m/N \to \lambda \in [0, 1]$ as $m \wedge n \to \infty$ where $N = m + n$, find the true asymptotic size of this test for non-normal X's and Y's as above.