Statistics 581, Problem Set 4 Solutions

Wellner; 10/28/2009

1. Ferguson, ACILST, problem 5, page 50: (The Poisson dispersion test). A standard test of the hypothesis H_0 that a distribution is Poisson(λ) for some λ is to reject H_0 if the ratio of the sample variance to the sample mean, $S_n^2/\overline{X}_n$, is too large. This test is good against alternatives whose variance is greater than the mean, such as the negative binomial distribution or any other mixture of Poisson distributions.
(a) Find the asymptotic distribution of $S_n^2/\overline{X}_n$ for general distributions.
(b) Find the asymptotic distribution of $S_n^2/\overline{X}_n$ under H_0 and show that it is independent of λ.

Solution: (a) We can use the result of part (a) of problem 3 of Problem Set #3. We just need to proceed as in (b) of problem 3, Problem Set #3 with $g(u,v) = v/u$. Thus we find that $\nabla g(u,v) = (-v/u^2, 1/u) = (-v/u, 1)/u$. Hence $\nabla g(\mu, \sigma^2) = (-\sigma^2/\mu, 1)/\mu$, and the limiting variance is

$$\begin{aligned} \nabla g^T \Sigma \nabla g &= \frac{\sigma^4}{\mu^2}\left(\frac{\sigma^2}{\mu^2} - 2\frac{\mu_3}{\mu\sigma^2} + \frac{\mu_4}{\sigma^4} - 1\right) \\ &= \frac{\sigma^4}{\mu^2}\left(\frac{\sigma^2}{\mu^2} - 2\frac{\sigma\gamma_1}{\mu} + 2 + \gamma_2\right) \equiv V^2 \,. \qquad (1) \end{aligned}$$

Thus it follows that

$$\sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma^2}{\mu}\right) \to_d N(0, V^2)$$

where V^2 is given in (1).
(b) When $X \sim$ Poisson(λ), $E(X) = \lambda$, $Var(X) = \lambda$, $\gamma_1 = 1/\sqrt{\lambda}$, and $\gamma_2 = 1/\lambda$. Thus we find that the asymptotic variance above is

$$\frac{\lambda^2}{\lambda^2}\left\{\frac{\lambda}{\lambda^2} - 2\frac{\lambda^{1/2}\lambda^{-1/2}}{\lambda} + 2 + \frac{1}{\lambda}\right\} = 2\,.$$

Thus it follows that under $X \sim$ Poisson(λ) we have

$$\sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma^2}{\mu}\right) = \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - 1\right) \to_d N(0, 2)\,.$$

2. (Continuation of problem 1 above.) Suppose that $(X|\Lambda) \sim$ Poisson(Λ) where $\Lambda \sim \Gamma(r, b)$ with density $b^r\lambda^{r-1}\exp(-b\lambda)/\Gamma(r)$ for some $r > 0$ and $b > 0$.
(a) Show that the marginal distribution of X is Negative Binomial $(b/(1+b), r)$ with density (probability mass function)

$$P_{r,b}(X = x) = \frac{\Gamma(r+x)}{x!\Gamma(r)}\left(\frac{b}{1+b}\right)^r \frac{1}{(1+b)^x}$$

for $x = 0, 1, \ldots$.
(b) Show that $E(X) = r/b$ and $Var(X) = (r/b) + r/b^2 > (r/b) = E(X)$, and hence if $b \equiv b_n = \sqrt{n}/\lambda_0$ and $r \equiv r_n = \sqrt{n}$, we have, letting E_n and Var_n denote expectation and variance under (r_n, b_n), $E_n(X) \to \lambda_0$ and $Var_n(X) \to \lambda_0$, while $\sqrt{n}(Var_n(X) - \lambda_0) \to \lambda_0^2$. (Hint: Use our results for computing the mean and variance conditionally on another random variable.)
(c) Show that if $X_n \sim$ Negative Binomial$(b_n/(1+b_n), r_n)$ with b_n and r_n as in (b), then $X_n \to_d X_0 \sim$ Poisson(λ_0).
(d) Now suppose that $X_{ni} \sim$ Negative Binomial$(b_n/(1+b_n), r_n)$ for $i = 1, \ldots, n$ are independent with b_n and r_n as in (b). Let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_{ni}$ and $S_n^2 = (n-1)^{-1}\sum_1^n (X_{ni} - \overline{X}_n)^2$ as in problem 1. Use the results of (b) to show that under this family of local alternatives to the Poisson distribution we have

$$\sqrt{n}(S_n^2/\overline{X}_n - 1) \to_d N(c, 2)$$

for some $c \neq 0$ and find c. Use this to approximate the power of the test in problem 1 for this particular sequence of alternatives.

Solution: (a) By direct calculation

$$\begin{aligned}
P_{r,b}(X = x) &= EP(X = x|\Lambda) \\
&= \int_0^\infty \frac{\lambda^x}{x!} e^{-\lambda} \frac{b(b\lambda)^{r-1}}{\Gamma(r)} \exp(-b\lambda) d\lambda \\
&= \frac{b^r}{x!\Gamma(r)} \int_0^\infty \lambda^{x+r-1} \exp(-(1+b)\lambda) d\lambda \\
&= \frac{b^r}{x!\Gamma(r)} \frac{1}{(1+b)^{r+x}} \int_0^\infty ((1+b)\lambda)^{r+x-1} e^{-(1+b)\lambda} (1+b) d\lambda \\
&= \frac{b^r}{x!\Gamma(r)} \frac{1}{(1+b)^{r+x}} \int_0^\infty v^{r+x-1} e^{-v} dv \\
&= \frac{\Gamma(r+x)}{x!\Gamma(r)} \left(\frac{b}{1+b}\right)^r \frac{1}{(1+b)^x}
\end{aligned}$$

for $x = 0, 1, 2, \ldots$. That is, $X \sim$ Negative Binomial$(p = b/(1+b), r)$.
(b) By computing conditionally on Λ we get

$$\begin{aligned}
&E(X) = E(E(X|\Lambda)) = E\Lambda = r/b, \\
&Var(X) = E(Var(X|\Lambda)) + Var(E(X|\Lambda)) \\
&\qquad = E(\Lambda) + Var(\Lambda) = r/b + r/b^2.
\end{aligned}$$

Thus when $b \equiv b_n = \sqrt{n}/\lambda_0$ and $r \equiv r_n = \sqrt{n}$, we have,

$$\begin{aligned}
&E_n(X) = r_n/b_n = \sqrt{n}/(\sqrt{n}/\lambda_0) = \lambda_0, \qquad \text{and} \\
&Var_n(X) = r_n/b_n + r_n/b_n^2 = \lambda_0 + \sqrt{n}/(n/\lambda_0^2) \\
&\qquad = \lambda_0 + \lambda_0^2/\sqrt{n} \to \lambda_0,
\end{aligned}$$

while

$$\sqrt{n}(Var_n(X) - \lambda_0) = \lambda_0^2.$$

(c) Since $E_n(\Lambda) = r_n/b_n = \lambda_0$ and

$$Var_n(\Lambda) = r_n/b_n^2 = \lambda_0^2/\sqrt{n} \to 0,$$

it follows that $\Lambda_n \to_p \lambda_0$; that is the sequence of mixing distributions converges to the distribution degenerate at λ_0. Hence for each $x \in \{0, 1, \ldots\}$ we have

$$\begin{aligned} P_{r_n,b_n}(X = x) &= EP(X = x|\Lambda_n) = E\left\{\frac{\Lambda_n^x}{x!}\exp(-\Lambda_n)\right\} \\ &\to \frac{\lambda_0^x}{x!}\exp(-\lambda_0) \qquad (2) \end{aligned}$$

by the Helly-Bray theorem applied to the bounded continuous function $g(v) = v^x \exp(-v)$ on $[0, \infty)$ and $\Lambda_n \to_d \lambda_0$. This implies that $X_n \to_d X_0 \sim Poisson(\lambda_0)$. The convergence in (2) can also be seen directly from the results in (a) as follows:

$$\begin{aligned} P_{r_n,b_n}(X = x) &= \frac{\Gamma(r_n + x)}{x!\Gamma(r_n)}\left(\frac{b_n}{1 + b_n}\right)^{r_n}\frac{1}{(1 + b_n)^x} \\ &= \frac{1}{x!}\frac{\Gamma(r_n + x)}{\Gamma(r_n)(1 + b_n)^x}\left(\frac{1}{1 + (1/b_n)}\right)^{r_n} \qquad (3) \end{aligned}$$

where

$$\left(\frac{1}{1 + (1/b_n)}\right)^{r_n} = \frac{1}{(1 + \frac{\lambda_0}{\sqrt{n}})^{\sqrt{n}}} \to \frac{1}{\exp(\lambda_0)} = \exp(-\lambda_0)$$

and, using Stirling's formula, $\Gamma(r + 1) \sim (\sqrt{2\pi r}(r/e)^r$ as $r \to \infty$,

$$\begin{aligned} \frac{\Gamma(r_n + x)}{\Gamma(r_n)(1 + b_n)^x} &\sim \frac{\sqrt{2\pi(r_n + x - 1)}}{\sqrt{2\pi(r_n - 1)}}\frac{\left(\frac{r_n+x-1}{e}\right)^{r_n+x-1}}{\left(\frac{r_n-1}{e}\right)^{r_n-1}}\frac{\lambda_0^x}{(\sqrt{n})^x} \\ &= \sqrt{\frac{r_n + x - 1}{r_n - 1}}\left(\frac{r_n + x - 1}{r_n - 1}\right)^{r_n-1} e^{-x}\left(\frac{r_n + x - 1}{\sqrt{n} - 1}\right)^x \cdot \lambda_0^x \\ &\to 1 \cdot e^x e^{-x} \cdot 1 \cdot \lambda_0^x = \lambda_0^x. \end{aligned}$$

Thus the expression on the right side in (3) converges to the Poisson probability

$$\frac{1}{x!}\lambda_0^x \exp(-\lambda_0).$$

(d) First write

$$\sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - 1\right) = \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma_n^2}{\mu_n}\right) + \sqrt{n}\left(\frac{\sigma_n^2}{\mu_n} - 1\right) \qquad (4)$$

where

$$\begin{aligned} \sigma_n^2 &= Var_n(X) = \lambda_0 + n^{-1/2}\lambda_0^2, \\ \mu_n &= E_n(X) = \lambda_0. \end{aligned}$$

Thus

$$\sqrt{n}\left(\frac{\sigma_n^2}{\mu_n} - 1\right) = \sqrt{n}(1 + n^{-1/2}\lambda_0 - 1) = \lambda_0, \qquad (5)$$

and if we show that

$$(6) \qquad \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma_n^2}{\mu_n}\right) \to_d N(0,2),$$

where $Z \sim N(0,1)$, then it follows from (4), (5), (6), that

$$T_n \equiv \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - 1\right) \to_d N(0,2) + \lambda_0 \sim N(\lambda_0, 2).$$

This yields an approximation to the power of the test derived in problem 2 under this sequence of Negative binomial alternatives:

$$\begin{aligned} \text{Power}(NegBin(r_n, b_n)) &= P_{r_n,b_n}(T_n > \sqrt{2}z_\alpha) \to P(\sqrt{2}Z + \lambda_0 > \sqrt{2}z_\alpha) \\ &= P(Z > z_\alpha - \lambda_0/\sqrt{2}) = 1 - \Phi(z_\alpha - \lambda_0/\sqrt{2}). \end{aligned}$$

(This is the power of an ad-hoc test based on comparison of two different moments of the Poisson distribution. What is the limiting power of an optimal test for distinguishing this sequence of negative binomial alternatives? We will return to this later!)

To show that (6) holds, suppose that we show that

$$(7) \qquad \sqrt{n}\begin{pmatrix} \overline{X}_n - \mu_n \\ S_n^2 - \sigma_n^2 \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \lambda_0 & \lambda_0 \\ \lambda_0 & \lambda_0 + 2\lambda_0^2 \end{pmatrix}.$$

Then $\overline{X}_n \to_p \lambda_0$ and

$$\begin{aligned} \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma_n^2}{\mu_n}\right) &= \sqrt{n}\frac{S_n^2 - \sigma_n^2}{\overline{X}_n} - \frac{\sigma_n^2}{\mu_n \overline{X}_n}\sqrt{n}(\overline{X}_n - \mu_n) \\ &\to_d \frac{Z_2}{\lambda_0} - \frac{\lambda_0}{\lambda_0^2}Z_1 = \lambda_0^{-1}(Z_2 - Z_1) \\ &\sim N(0,2). \end{aligned}$$

Another way to arrange the argument once (7) is proved is to note that (7) implies that

$$(8) \qquad \sqrt{n}\begin{pmatrix} \overline{X}_n - \lambda_0 \\ S_n^2 - \lambda_0 \end{pmatrix} \to_d \underline{Z} + \begin{pmatrix} 0 \\ \lambda_0^2 \end{pmatrix} \sim N_2((0, \lambda_0^2)^T, \Sigma).$$

Then we can apply the delta-method directly with $g(u,v) = v/u$ as before: since $g'(\lambda_0, \lambda_0) = (-1, 1)/\lambda_0$,

$$\begin{aligned} \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - 1\right) &= \sqrt{n}(g(\overline{X}_n, S_n^2) - g(\lambda_0, \lambda_0)) \\ &\to_d g'(\underline{Z} + (0, \lambda_0^2)^T) \sim N(0,2) + \lambda_0 = N(\lambda_0, 2). \end{aligned}$$

The rest of the proof is concerned (only) with showing that (7) holds. As before,

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu_n \\ S_n^2 - \sigma_n^2 \end{pmatrix} = \sqrt{n}\begin{pmatrix} \overline{X}_n - \mu_n \\ \overline{Y}_n \end{pmatrix} + o_p(1) \equiv Z_n + o_p(1)$$

where $Y_i = (X_i - \mu_n)^2 - \sigma_n^2$, so it suffices to show that $Z_n \to_d N_2(0, \Sigma)$.

To do this, let $a \in \mathbb{R}^2$, and consider

$$\begin{aligned} a'Z_n &= a_1\sqrt{n}(\overline{X}_n - \mu_n) + a_2\sqrt{n}\overline{Y}_n \\ &= \sum_{i=1}^n \{a_1 n^{-1/2}(X_i - \mu_n) + a_2 n^{-1/2}[(X_i - \mu_n)^2 - \sigma_n^2]\} \\ &= \sum_{i=1}^n X_{ni} \end{aligned}$$

where $\mu_{ni} = EX_{ni} = 0$ and

$$\sigma_{ni}^2 = Var(X_{ni}) = a_1^2 n^{-1} Var_n(X_1) + 2a_1a_2n^{-1}E_n(X_1 - \mu_n)^3 + a_2^2 n^{-1}\{E_n(X_1 - \mu_n)^4 - \sigma_n^4\},$$

so that $\mu_n = \sum_1^n \mu_{ni} = 0$, and

$$\begin{aligned} \sigma_n^2 &= \sum_{i=1}^n \sigma_{ni}^2 = a_1^2 Var_n(X_1) + 2a_1a_2E_n(X_1 - \mu_n)^3 + a_2^2\{E_n(X_1 - \mu_n)^4 - \sigma_n^4\} \\ &\to a'\Sigma a > 0 \end{aligned} \tag{9}$$

since, using the results of (b)

$$\begin{aligned} &Var_n(X) = \lambda_0 + \lambda_0^2/\sqrt{n} \to \lambda_0, \\ &E_n(X_1 - \mu_n)^3 = E_n E[(X_1 - \mu_n)^3 | \Lambda_n] \to \lambda_0, \\ &E_n(X_1 - \mu_n)^4 - \sigma_n^4 \to \lambda_0 + 2\lambda_0^2. \end{aligned}$$

Now we verify the hypothesis of the Liapunov CLT: first,

$$\begin{aligned} \gamma_{ni} &\equiv E|X_{ni}|^3 \\ &\le 2^2\{E_n|a_1 n^{-1/2}(X_i - \mu_n)|^3 + E_n|a_2 n^{-1/2} Y_i|^3\} \\ &\le 2^2\{|a_1|^3 n^{-3/2} E_n|X_1 - \mu_n|^3 + |a_2|^3 n^{-3/2}|Y_1|^3\}. \end{aligned}$$

Therefore

$$\begin{aligned} \gamma_n &\equiv \sum_{i=1}^n \gamma_{ni} \\ &\le 2^2\{|a_1|^3 n^{-1/2} E_n|X_1 - \mu_n|^3 + |a_2|^3 n^{-1/2} E_n|Y_i|^3\}. \end{aligned}$$

Thus we see that $\gamma_n/\sigma_n^3 \to 0$ if both $\limsup_{n\to\infty} E_n|X_1 - \mu_n|^3 < \infty$ and $\limsup_{n\to\infty} E_n|(X_1 - \mu_n)^2 - \sigma_n^2|^3 < \infty$. By applying Minkowski's inequality (the triangle inequality for the L_3 norm), it is clear that this will hold if the same is true for $E_n|X_1|^3$ and $E_n|X_1|^6$. But note that for $x \in \{0, 1, \ldots\}$, any positive integer m and number $t > 0$, $t^m|x|^m/m! = (tx)^m/m! \le \exp(tx)$, so $E_n|X_1|^m \le (m!/t^m)E_n\exp(tX_1)$, and hence it suffices to show that $\limsup_{n\to\infty} E_n\exp(tX_1) < \infty$. But since the moment generating function of $U \sim Poisson(\lambda)$ is $E\exp(tU) = \exp(\lambda(e^t-1))$

$$\begin{aligned}
E_n\exp(tX_1) &= E_nE[\exp(tX_1)|\Lambda_n] = E_n\exp(\Lambda_n(e^t-1)) \\
&= \int_0^\infty \exp(\lambda(e^t-1))\frac{b_n(b_n\lambda)^{r_n-1}}{\Gamma(r_n)}\exp(-b_n\lambda)d\lambda \\
&= \int_0^\infty \frac{b_n(b_n\lambda)^{r_n-1}}{\Gamma(r_n)}\exp(-(b_n-(e^t-1))\lambda)d\lambda \\
&= \left(\frac{b_n}{b_n-(e^t-1)}\right)^{r_n} \\
&= \left(1-\frac{e^t-1}{b_n}\right)^{-r_n} \\
&= \left(1-\frac{\lambda_0(e^t-1)}{\sqrt{n}}\right)^{-\sqrt{n}} \\
&\to \exp(\lambda_0(e^t-1)).
\end{aligned}$$

(Note that this is, in fact, the moment generating function of a Poisson random variable with parameter λ_0.) Thus $\gamma_n/\sigma_n^3 \to 0$ holds, and via the Liapunov CLT we find that $a'Z_n/\sigma_n \to_d N(0,1)$. In view of (9) this yields $a'Z_n \to_d a'Z \sim N(0, a'\Sigma a)$, and by the Cramér-Wold device we conclude $Z_n \to_d Z \sim N_2(0,\Sigma)$.

3. 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 $\underline{p}_n$.)

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

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

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

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

$$\equiv \sum_{i=1}^{n} X_{ni}$$

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

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

where

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

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

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

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

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

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 (11) 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,

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

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

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

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 4. 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 $n^{-1}Q_n \rightarrow_p$ 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 $n = 50$, $\alpha = .05$, 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

$$\begin{aligned} Q_n &= \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{(N_{ij} - n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}} \\ &= \frac{(N_{11}N_{22} - N_{12}N_{21})^2}{n^3} \sum_{i,j} \left\{ \frac{1}{\hat{p}_{i\cdot}\hat{p}_{\cdot j}} \right\} \\ &= \frac{(N_{12}N_{21} - N_{11}N_{22})^2}{n^3} \frac{1}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\ &= \frac{n\{\exp(\hat{\psi}_n) - 1\}^2\,(\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\ &= \frac{\{\sqrt{n}[\exp(\hat{\psi}_n) - 1]\}^2\,(\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})} \\ &\to_d [N(0,V^2)]^2 \frac{[p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1})]^2}{p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1})} \\ &= [N(0,V^2)]^2 p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1}) = [N(0,1)]^2 \stackrel{d}{=} \chi_1^2 \end{aligned}$$

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

$$Q_n = \frac{n\{\exp(\hat{\psi}_n) - 1\}^2\,(\hat{p}_{11}\hat{p}_{22})^2}{\hat{p}_{1\cdot}(1-\hat{p}_{1\cdot})\hat{p}_{\cdot 1}(1-\hat{p}_{\cdot 1})}$$

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

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 $n = 50$, $\alpha = .05$, 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 $c = p_{1\cdot}(1-p_{1\cdot})p_{\cdot 1}(1-p_{\cdot 1}) = (.5)^2(.4)(.6) = .06$),

$$
t_n \equiv \sqrt{n} \log \frac{p_{12}p_{21}}{p_{11}p_{22}} = -12.67... \, .
$$

Thus $\delta = (12.67\ldots)^2(.06) = 9.631...$, and an approximation to the power of our test is given by

$$
P(\chi_1^2(9.631...) > \chi^2_{1,.05}) = P(\chi_1^2(9.631...) > 3.841...) = .874... \, .
$$

6. Suppose that $X_{n,1}, \ldots, X_{n,n}$ are independent Bernoulli$(p_{n,1}), \ldots,$ Bernoulli$(p_{n,n})$ respectively. Let $T_n = X_{n,1} + \cdots + X_{n,n}$, $\mu_n = \sum_{i=1}^n p_{n,i}$, and $\sigma_n^2 \equiv \sum_{i=1}^n p_{n,i}(1-p_{n,i})$.
(a) Use the Liapunov central limit theorem to show that if $\sigma_n^2 \to \infty$, then $(T_n - \mu_n)/\sigma_n \to_d N(0,1)$.
(b) Show that the key condition of the Liapunov CLT implies the Lindeberg condition.

Solution: (a) We compute $\mu_{ni} \equiv EX_{ni} = p_{ni}$, $\sigma_{ni}^2 = Var(X_{ni}) = p_{ni}(1-p_{ni})$, and

$$
\begin{aligned}
\gamma_{ni} &\equiv E|X_{ni} - p_{ni}|^3 = p_{ni}|1-p_{ni}|^3 + (1-p_{ni})|0 - p_{ni}|^3 \\
&= p_{ni}(1-p_{ni})\{(1-p_{ni})^2 + p_{ni}^2\} \\
&\le p_{ni}(1-p_{ni}) = \sigma_{ni}^2.
\end{aligned}
$$

Thus $\mu_n \equiv \sum_1^n \mu_{ni}$, $\sigma_n^2 = \sum_1^n p_{ni}(1-p_{ni})$, and

$$
\gamma_n \equiv \sum_1^n \gamma_{ni} \le \sum_1^n \sigma_{ni}^2 = \sigma_n^2.
$$

It follows that $\gamma_n/\sigma_n^3 \le \sigma_n^2/\sigma_n^3 = 1/\sigma_n \to 0$ if $\sigma_n^2 \to \infty$. So the Liapunov CLT yields $(T_n - \mu_n)/\sigma_n \to_d N(0,1)$ under the condition $\sigma_n^2 = \sum_1^n p_{ni}(1-p_{ni}) \to \infty$.

(b) Note that

$$
|w|^r 1_{[|w|>c]} \le |w|^r \frac{|w|}{c} 1_{[|w|>c]} \le \frac{|w|^{r+1}}{c}.
$$

Using this with $r = 2$, $c = \epsilon\sigma_n$, $w = X_{ni}$ yields

$$
\begin{aligned}
\frac{1}{\sigma_n^2}\sum_{i=1}^n E\{|X_{ni}|^2 1_{[|X_{ni}|>\epsilon\sigma_n]}\} &\le \frac{1}{\sigma_n^2}\sum_{i=1}^n E\{|X_{ni}|^2 \frac{|X_{ni}|}{\epsilon\sigma_n} 1_{[|X_{ni}|>\epsilon\sigma_n]}\} \\
&\le \frac{1}{\sigma_n^2}\sum_{i=1}^n \frac{E\{|X_{ni}|^3\}}{\epsilon\sigma_n} \\
&= \frac{1}{\epsilon\sigma_n^3}\sum_{i=1}^n E|X_{ni}|^3 = \frac{1}{\epsilon}\frac{\gamma_n}{\sigma_n^3}
\end{aligned}
$$

for X_{ni}'s with $EX_{ni} = 0$. Thus $\gamma_n/\sigma_n^3 \to 0$ implies that the Lindeberg condition holds.