**Statistics 581, Problem Set 4 Solutions**

Wellner; 10/27/2017

1. 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) Suppose that $X_{n,i}$ and $T_n$ are as in (a) with $p_{n,i} = p_n \equiv p_0 + cn^{-1/2}$ for $1 \le i \le n$ and $n \ge c^2/\min\{p_0^2, (1-p_0)^2\}$. Show that $\sqrt{n}(\widehat{p}_n - p_n) \to_d N(0, p_0(1-p_0))$.
(c) Show that the key condition of the Liapunov CLT implies the Lindeberg condition (and hence the Lindeberg-Feller CLT also holds).

**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) When $p_{ni} = p_0 + cn^{-1/2}$,

$$\begin{aligned}
&\sum_1^n p_{n,i} = np_n = np_0 + cn^{1/2}, \quad \text{and} \quad n^{-1}\sum_1^n p_{n,i} = p_n \to p_0, \\
&\sigma_n^2 = \sum_{i=1}^n p_{n,i}(1-p_{n,i}) = np_n(1-p_n) \to \infty, \\
&\frac{\sigma_n^2}{n} = p_n(1-p_n) \to p_0(1-p_0).
\end{aligned}$$

Thus the condition $\sigma_n^2 \to \infty$ in (a) holds and we have $(T_n - np_n)/\sigma_n \to_d Z \sim N(0,1)$. But this yields

$$\sqrt{n}(\hat{p}_n - p_n) = \frac{\sigma_n}{\sqrt{n}}\frac{T_n - np_n}{\sigma_n} \to_d \sqrt{p_0(1-p_0)}Z \sim N(0, p_0(1-p_0)).$$

(c) 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.

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 $H_0 : \underline{p}_0 = (1/3, 1/3, 1/3)$, 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 $\underline{N}_n = \sum_{i=1}^n \underline{\Delta}_{ni}$ where $\underline{\Delta}_{ni} \sim \text{Mult}_k(1, \underline{p}_n)$ are i.i.d. for each $n$, we can write

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

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

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

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

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

with $\widetilde{\Sigma} \equiv \text{diag}(p_0) - \underline{p}_0\underline{p}_0^T$, 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 $\sigma_n^2 \to \underline{a}^T\widetilde{\Sigma}\underline{a}$, this implies

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

and by Cramér - Wold, this yields

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

(b) Now

$$\begin{aligned}n^{1/2}(\underline{p} - \underline{p}_0) &= 10((.2,.6,.2) - (1/3,1/3,1/3)) = 10(-2/15, 4/15, -2/15) \\ &= 10(-2,4,-2)/15 = (-4,8,-4)/3,\end{aligned}$$

so the non-centrality parameter is

$$\delta = 9^{-1}\left\{\frac{(4)^2}{1/3} + \frac{8^2}{1/3} + \frac{(4)^2}{1/3}\right\} = 96/3 = 32.$$

Thus the approximate power via $\chi^2_2(\delta)$ is

$$P(\chi^2_2(32) \ge \chi_{2,.05}) = P(\chi^2_2(32) \ge 5.991) = 0.9996\ldots\,, \qquad \text{when} \quad \alpha = .05\,,$$

and

$$P(\chi^2_2(4) \ge \chi_{2,.01}) = P(\chi^2_2(4) \ge 9.210) = 0.997\ldots \qquad \text{when} \quad \alpha = .01\,,$$

Now we want to find $n$ so that

$$P(\chi^2_2(\delta_n) \ge 5.991) = .90$$

where

$$\delta_n = n\left\{\frac{(2/15)^2}{1/3} + \frac{(4/15)^2}{1/3} + \frac{(2/15)^2}{1/3}\right) = n(32/100)$$

In this case we find that $\delta_n = n(32/100) = 12.6539$, so that $n = (100/32)\cdot 12.6539 \approx 40$. When $\alpha = .01$ we find that $\delta_n = n(32/100) = 17.4267$ so that $n = (100/32)\cdot 17.4267/(.04) \approx 54$.

3. Ferguson, ACILST, problem 5, page 50: (The Poisson dispersion test). A standard test of the hypothesis $H_0$ that a distribution is Poisson($\lambda$) for some $\lambda$ is to reject $H_0$ if the ratio of the sample variance to the sample mean, $S^2_n/\overline{X}_n$, 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 $S^2_n/\overline{X}_n$ for general i.i.d. random variables $X_1, \ldots, X_n$ with $EX_1 > 0$ and $E|X_1|^4 < \infty$; i.e. show that $\sqrt{n}(S^2_n/\overline{X}_n - \sigma^2/\mu) \to_d$ "something" and find "something".
(b) Find the asymptotic distribution of $S^2_n/\overline{X}_n$ under $H_0$ and show that it is independent of $\lambda$.

**Solution:** (a) We can use the result of part (a) of problem 2 of Problem Set #3. We just need to proceed as in (b) of problem 2, 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

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

Thus it follows that

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

where $V^2$ is given in (2).
(b) When $X \sim$ Poisson($\lambda$), $E(X) = \lambda$, $Var(X) = \lambda$, $\gamma_1 = 1/\sqrt{\lambda}$, 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 $X \sim \text{Poisson}(\lambda)$ we have

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

4. (Continuation of problem 3 above.) Suppose that $(X|\Lambda) \sim \text{Poisson}(\Lambda)$ 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 $(r_n, b_n)$, $E_n(X) \rightarrow \lambda_0$ and $Var_n(X) \rightarrow \lambda_0$, while $\sqrt{n}(Var_n(X) - \lambda_0) \rightarrow \lambda_0^2$. (Hint: Use our results for computing the mean and variance conditionally on another random variable.)
(c) Show that if $X_n \sim \text{Negative Binomial}(b_n/(1+b_n), r_n)$ with $b_n$ and $r_n$ as in (b), then $X_n \rightarrow_d X_0 \sim \text{Poisson}(\lambda_0)$.
(d) Now suppose that $X_{ni} \sim \text{Negative Binomial}(b_n/(1+b_n), r_n)$ 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

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

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, $X \sim \text{Negative Binomial}(p = b/(1+b), r)$.
(b) By computing conditionally on $\Lambda$ we get

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

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

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

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 $\lambda_0$. Hence for each $x \in \{0, 1, \ldots\}$ we have

$$\begin{aligned} P_{r_n,b_n}(X = x) &= EP(X = x|\Lambda_n) = E\left\{\frac{\Lambda_n^x}{x!}\exp(-\Lambda_n)\right\} \\ &\to \frac{\lambda_0^x}{x!}\exp(-\lambda_0) \end{aligned} \tag{3}$$

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 (3) can also be seen directly from the results in (a) as follows:

$$\begin{aligned} P_{r_n,b_n}(X = x) &= \frac{\Gamma(r_n + x)}{x!\Gamma(r_n)}\left(\frac{b_n}{1 + b_n}\right)^{r_n}\frac{1}{(1 + b_n)^x} \\ &= \frac{1}{x!}\frac{\Gamma(r_n + x)}{\Gamma(r_n)(1 + b_n)^x}\left(\frac{1}{1 + (1/b_n)}\right)^{r_n} \end{aligned} \tag{4}$$

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 (4) converges to the Poisson probability

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

(d) First write

$$\sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - 1\right) = \sqrt{n}\left(\frac{S_n^2}{\overline{X}_n} - \frac{\sigma_n^2}{\mu_n}\right) + \sqrt{n}\left(\frac{\sigma_n^2}{\mu_n} - 1\right) \tag{5}$$

where

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

Thus

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

and if we show that

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

where $Z \sim N(0,1)$, then it follows from (5), (6), (7), that

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

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

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

(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 (7) holds, suppose that we show that

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

where

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

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

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

Another way to arrange the argument once (8) is proved is to note that (8) implies that

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

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 (8) holds. As before,

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

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

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

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

$$\begin{aligned}\sigma_n^2 &= \sum_{i=1}^n \sigma_{ni}^2 = a_1^2 Var_n(X_1) + 2a_1 a_2 E_n(X_1 - \mu_n)^3 + a_2^2\{E_n(X_1 - \mu_n)^4 - \sigma_n^4\} \\ (10) \quad &\to a'\Sigma a > 0\end{aligned}$$

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 $\lambda_0$.) 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 (10) this yields $a'Z_n \to_d a'Z \sim N(0, a'\Sigma a)$, and by the Cramér-Wold device we conclude $Z_n \to_d Z \sim N_2(0,\Sigma)$.