Statistics 581, Problem Set 4 Solutions

Wellner; 11/2/2001

1. Suppose that $\underline{N}_n \sim \text{Mult}_k(n, \underline{p})$ and $\widehat{\underline{p}} = \underline{N}_n/n$. Suppose that $g : R^k \to R^k$ is of the form $g(\underline{x}) = (g_1(x_1), \ldots, g_k(x_k))$ where each g_j is differentiable. Then the "transformed chi-square statistic" $C_n(g)$ is defined by

$$C_n(g) \equiv C_n(g, \underline{p}) = n \sum_{j=1}^{k} \frac{(g_j(\widehat{p}_j) - g_j(p_j))^2}{p_j \dot{g}_j(p_j)^2} .$$

(a) Show that $C_n(g) \to_d \chi^2_{k-1}$.
(b) Specialize this to the case $g_j(x_j) = x_j^{1/2}$, and show that the resulting statistic is related to the Hellinger distance between $\widehat{\underline{p}}$ and $\underline{p}$.
(c) Suppose that the "true $\underline{p}$ is $\underline{p}_n = \underline{p}_0 + n^{-1/2}\underline{c}$. Thus $\underline{N}_n = \sum_{i=1}^n \underline{M}_{ni}$ where $\underline{M}_{n1}, \ldots, \underline{M}_{n1}$ are i.i.d. $\text{Mult}_k(1, \underline{p}_n)$.
Show that $C_n(g, \underline{p}_0) \to_d \chi^2_{k-1}(\delta)$ where $\delta = \sum_{j=1}^k c_j^2/p_{0j}$.
[Hint: See Ferguson pages 59 and 66.]
Solution: (a) We know that

$$n^{1/2}(\widehat{\underline{p}}_n - \underline{p}) \to_d N_k(0, A)$$

where $A = \text{diag}(\underline{p}) - \underline{p}\underline{p}'$. Hence, by the delta-method,

$$n^{1/2}(\underline{g}(\widehat{\underline{p}}_n) - \underline{g}(\underline{p})) \to_d \dot{g} N_k(0, A) = N_k(0, \dot{g} A \dot{g}')$$

where $\dot{g} \equiv \text{diag}(\dot{\underline{g}})$. Hence

$$\underline{Z}_n(g) \equiv \dot{g}^{-1}\text{diag}(1/\sqrt{\underline{p}}) n^{1/2}(\underline{g}(\widehat{\underline{p}}_n) - \underline{g}(\underline{p})) \to_d \underline{Z} \sim N_k(0, \Sigma)$$

where $\Sigma = I - \sqrt{\underline{p}}\sqrt{\underline{p}}'$. Since

$$C_n(g, \underline{p}) = \|\underline{Z}_n(g)\|^2 ,$$

it follows from the continuous mapping theorem that

$$C_n(g, \underline{p}) \to_d \|\underline{Z}\|^2$$

just as in the untransformed case treated in class, and the distribution of the random variable on the right side is just χ^2_{k-1}.
(b) When $g_j(x_j) = x_j^{1/2}$ for $j = 1, \ldots, k$, then $\dot{g}_j(p_j) = (1/2)p_j^{-1/2}$, and hence

$$\dot{g}^{-1}\text{diag}(1/\sqrt{\underline{p}}) n^{1/2}(\underline{g}(\widehat{\underline{p}}) - \underline{g}(\underline{p})) = 2n^{1/2}(\sqrt{\widehat{\underline{p}}} - \sqrt{\underline{p}})$$

and

$$C_n(g) = 4n \sum_{j=1}^{k} (\sqrt{\widehat{p}_j} - \sqrt{p_j})^2 = 4n d_H^2(\underline{\widehat{p}}, \underline{p})$$

where $d_H(\underline{p}, \underline{q})$ is the Hellinger distance between $\underline{p}$ and $\underline{q}$.

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

$$\underline{Z}_n \equiv \text{diag}(1/\sqrt{\underline{p}_0}) n^{1/2} (\underline{\widehat{p}} - \underline{p}) \to \underline{Z} + \underline{d} \sim N_k(\underline{d}, \Sigma) \tag{0.1}$$

where $\underline{d} = \text{diag}(1/\sqrt{\underline{p}_0})\underline{c}$ and $\Sigma = I - \sqrt{\underline{p}_0}\sqrt{\underline{p}_0}'$. If we accept (0.1), then it follows from the delta method that

$$n^{1/2}(\underline{g}(\underline{\widehat{p}}_n) - \underline{g}(\underline{p}_0)) \to_d \dot{g} \text{diag}(\sqrt{\underline{p}_0})(\underline{Z} + \underline{d})$$

where now $\dot{g} = \dot{g}(\underline{p}_0)$. This implies that

$$\underline{Z}_n(g) \equiv \dot{g}^{-1} \text{diag}(1/\sqrt{\underline{p}_0}) n^{1/2} (\underline{g}(\underline{\widehat{p}}_n) - \underline{g}(\underline{p}_0)) \to_d \underline{Z} + \underline{d}\,,$$

and hence by the continuous mapping theorem that

$$C_n(g, \underline{p}_0) \to_d \|\underline{Z} + \underline{d}\|^2$$

just as in the untransformed case. From the discussion in class we know that the distribution of the random variable on the right side is just $\chi^2_{k-1}(\delta)$ with $\delta = \sum_{j=1}^k c_j^2/p_{0j}$.

To prove that (0.1) holds, we can use the Cramér-Wold device and the Liapunov CLT. Fix $\underline{a} \in R^k$. Then we want to show that

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

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

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

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

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

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

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

while

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

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 $\gamma_n/\sigma_n^{3/2} \rightarrow 0$, and

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

This implies

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

and by Cramér - Wold, this implies

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

2. Ferguson, ACILST, problem 1, page 65: (a) In a multinomial experiment with sample size 100 and 3 cells with null hypothesis $H_0 : p_1 = 1/4, p_2 = 1/2, p_3 = 1/4$, what is the approximate power at the alternative $p_1 = 0.2, p_2 = 0.6, p_3 = 0.2$ when the level of significance is $\alpha = 0.05$? $\alpha = 0.01$? (b) How large a sample size is needed to achieve power 0.9 at this alternative when $\alpha = 0.05$? $\alpha = 0.01$?
Solution: Now

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

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

$$P(\chi_2^2(4) \geq \chi_{2,.05}) = P(\chi_2^2(4) \geq 5.991) = .415\,, \qquad \text{when} \;\; \alpha = .05\,,$$

and

$$P(\chi_2^2(4) \geq \chi_{2,.01}) = P(\chi_2^2(4) \geq 9.210) = .204 \qquad \text{when} \;\; \alpha = .01\,,$$

(b) Now we want to find n so that

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

where

$$\delta_n = n\left(\frac{.05^2}{.25} + \frac{.1^2}{.5} + \frac{.05^2}{.25}\right) = .04n\,.$$

In this case we find that $\delta_n = .04n = 12.6539$, so that $n = 12.6539/(.04) \approx 317$. When $\alpha = .01$ we find that $\delta_n = .04n = 17.4267$ so that $n = 17.4267/(.04) = 436$.

3. Ferguson, ACILST, problem 3, page 42: consider the autoregressive scheme $X_n = \beta X_{n-1} + \epsilon_n$, for $n = 1, 2, \ldots$, where $\epsilon_1, \epsilon_2, \ldots$ are i.i.d., $E(\epsilon_n = \mu$, $Var(\epsilon_n) = \sigma^2$, $-1 \leq \beta < 1$, and $X_0 = 0$. Show that $\overline{X}_n$ is asymptotically normal:

$$\sqrt{n}(\overline{X}_n - \mu/(1-\beta)) \to_d N(0, \sigma^2/(1-\beta)^2), \qquad \text{if} \qquad -1 < \beta < 1\,,$$

$$\sqrt{n}(\overline{X}_n - \mu/2) \to_d N(0, \sigma^2/2), \qquad \text{if} \qquad \beta = -1\,.$$

Proof. Note that

$$\begin{aligned}
X_1 + X_2 &= \epsilon_1 + \beta\epsilon_1 + \epsilon_2 = (1+\beta)\epsilon_1 + \epsilon_2, \\
X_1 + X_2 + X_3 &= (1+\beta)\epsilon_1 + \epsilon_2 + \beta^2\epsilon_1 + \beta\epsilon_2 + \epsilon_3\,, \\
X_1 + \ldots + X_n &= (1 + \beta + \ldots + \beta^{n-1})\epsilon_1 + (1 + \beta + \ldots + \beta^{n-2})\epsilon_2 + \ldots + \epsilon_3 \\
&= \sum_{i=1}^n a_{n,i}\epsilon_i \equiv \sum_{i=1}^n Y_{n,i}
\end{aligned}$$

where

$$a_{n,i} = \sum_{j=0}^{n-i} \beta^j, \quad i = 1, \ldots, n\,.$$

Now $\mu_{n,i} \equiv E(Y_{n,i}) = a_{n,i}E(\epsilon_i) = a_{n,i}\mu$ so that

$$\begin{aligned}
\mu_n \equiv E(\overline{X}_n) &= \frac{1}{n}\sum_{i=1}^n a_{n,i} = \frac{\mu}{n}\sum_{i=1}^n \sum_{j=0}^{n-i} \beta^j \\
&= \frac{\mu}{n}\sum_{j=0}^{\infty}\sum_{i=1}^n 1_{[j \leq n-i]}\beta^j = \frac{\mu}{n}\sum_{j=0}^{\infty}(n-j)\beta^j \\
&= \mu\left\{\sum_{j=0}^n \beta^j - n^{-1}\sum_{j=0}^n j\beta^j\right\} \to \mu\sum_{j=1}^{\infty}\frac{\mu}{1-\beta}\,.
\end{aligned}$$

Also,

$$\begin{aligned}
\sigma_n^2 \equiv Var(\sqrt{n}(\overline{X}_n - E(\overline{X}_n))) &= \frac{1}{n}\sum_{i=1}^n a_{n,i}^2\sigma^2 = \frac{\sigma^2}{n}\sum_{i=1}^n \left(\sum_{j=0}^{n-i}\beta^j\right)^2 \\
&= \frac{\sigma^2}{n}\sum_{j=0}^n\sum_{j'=0}^n\sum_{i=1}^n 1_{[j\le n-i, j'\le n-i]}\beta^{j+j'} \\
&= \frac{\sigma^2}{n}\sum_{j=0}^n\sum_{j'=0}^n (n-j)\wedge(n-j')\beta^{j+j'} \\
&= \sigma^2\sum_{j=0}^n\sum_{j'=0}^n (1-j/n)\wedge(1-j'/n)\beta^{j+j'} \\
&\to \sigma^2\sum_{j=0}^\infty\sum_{j'=0}^\infty \beta^{j+j'} \qquad \text{by the DCT} \\
&= \sigma^2\left(\sum_{j=0}^\infty \beta^j\right)^2 = \frac{\sigma^2}{(1-\beta)^2}\,.
\end{aligned}$$

To apply the Lindeberg-Feller CLT, we first write

$$\sqrt{n}(\overline{X}_n - E(\overline{X}_n)) = \sum_{i=1}^n b_{n,i}\tilde{\epsilon}_i$$

where $\tilde{\epsilon}_i \equiv \epsilon_i - \mu$ are i.i.d. with mean zero and variance σ^2, and $b_{n,i} = a_{n,i}/\sqrt{n}$. Thus the random variables of the Lindeberg-Feller CLT are $X_{n,i} = b_{n,i}\tilde{\epsilon}_i$. The Lindeberg-condition becomes

$$\begin{aligned}
&\frac{1}{\sigma_n^2}\sum_{i=1}^n E\{X_{n,i}^2 1_{[|X_{n,i}|>\epsilon\sigma_n]}\} \\
&\quad = \frac{1}{\sigma_n^2}\sum_{i=1}^n E\{b_{n,i}^2 E\{\tilde{\epsilon}_i^2 1_{[|\tilde{\epsilon}_i|>\epsilon\sigma_n/|b_{n,i}|]}\} \\
&\quad \le \frac{1}{\sigma_n^2}\sum_{i=1}^n b_{n,i}^2 E\{\tilde{\epsilon}_i^2 1_{[|\tilde{\epsilon}_i|>\epsilon\sigma_n/\max_{1\le i\le n}|b_{n,i}|]}\} \\
&\quad \le \frac{1}{\sigma_n^2}\frac{1}{n}\sum_{i=1}^n a_{n,i}^2 E\{\tilde{\epsilon}_i^2 1_{[|\tilde{\epsilon}_i|>\epsilon\sigma_n/\max_{1\le i\le n}|b_{n,i}|]}\} \\
&\quad = \frac{1}{\sigma^2}E\{\tilde{\epsilon}_1^2 1_{[|\tilde{\epsilon}_1|>\epsilon\sigma_n/\max_{1\le i\le n}|b_{n,i}|]}\} \\
&\quad \to 0 \qquad \text{for every } \epsilon
\end{aligned}$$

since $\sigma_n^2 \to \sigma^2/(1-\beta)^2$ as calculated above, and the DCT in view of the fact that

$$\max_{1\le i\le n} |b_{n,i}| \le \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} |\beta|^j \to 0\,.$$

Thus the Lindeberg-Feller CLT implies that

$$\sqrt{n}(\overline{X}_n - E(\overline{X}_n)) \to_d N\left(0, \frac{\sigma^2}{(1-\beta)^2}\right)$$

if $\beta \in (-1,1)$. Also note that

$$\sqrt{n}(E(\overline{X}_n) - \frac{\mu}{1-\beta}) = -\frac{1}{\sqrt{n}} \sum_{j=0}^{n} j\beta^j \to 0\,,$$

so we conclude that

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

When $\beta = -1$, $\overline{X}_n = (\epsilon_1 + \epsilon_3 + \ldots + \epsilon_n/n$, n odd, $\overline{X}_n = (\epsilon_2 + \epsilon_4 + \ldots + \epsilon_n)/n$, n even, so $E(\overline{X}_n) = ((n+1)/(2n))\mu$, n odd, $E(\overline{X}_n) = (1/2)\mu$, n even. Thus we have

$$\sqrt{(n+1)/2}(\overline{X}_n - \mu/2) \to_d N(0, \sigma^2)$$

as $n \to \infty$ through $n = 2m+1$, $m = 1, 2, 3, \ldots$, and

$$\sqrt{n/2}(\overline{X}_n - \mu/2) \to_d N(0, \sigma^2)$$

as $n \to \infty$ through $n = 2m$, $m = 1, 2, 3, \ldots$. Hence it follows that

$$\sqrt{n}(\overline{X}_n - \mu/2) \to_d N(0, 2\sigma^2)\,.$$

[Note the slight wobble in Ferguson's solution here!]
When $\beta = 1$, $a_{n,i} = n - i + 1$, and hence $n\overline{X}_n = \sum_{i=1}^n (n-i+1)\epsilon_i =_d \sum_{j=1}^n j\epsilon_j$. Thus

$$E(\overline{X}_n) = \mu \sum_{i=1}^n i/n = \mu \frac{n(n+1)}{2n} = \mu \frac{n+1}{2}\,,$$

and

$$Var(\overline{X}_n) = \frac{\sigma^2}{n^2} \sum_{i=1}^n i^2 = \frac{\sigma^2}{n^2} \frac{n(n+1)(2n+1)}{6} \sim \frac{n}{3}\sigma^2\,.$$

Now the Lindeberg condition becomes

$$\begin{aligned}
&\frac{1}{\sigma_n^2}\sum_{i=1}^n E\{i^2\tilde{\epsilon}_i^2 1_{[i|\tilde{\epsilon}_i|>\epsilon\sigma_n]}\} \\
&\quad = \frac{1}{\sigma_n^2}\sum_{i=1}^n E\{i^2\tilde{\epsilon}_i^2 1_{[i|\tilde{\epsilon}_i|>\epsilon\sigma_n]}\} \\
&\quad \le \frac{1}{\sigma_n^2}\sum_{i=1}^n i^2 E\{\tilde{\epsilon}_i^2 1_{[|\tilde{\epsilon}_i|>\epsilon\sigma_n/\max_{1\le i\le n} i]}\} \\
&\quad \le \frac{1}{\sigma_n^2}\sum_{i=1}^n i^2 E\{\tilde{\epsilon}_i^2 1_{[|\tilde{\epsilon}_i|>\epsilon\sigma_n/\max_{1\le i\le n} i]}\} \\
&\quad = \frac{1}{\sigma_n^2}\sum_{i=1}^n i^2 E\{\tilde{\epsilon}_1^2 1_{[|\tilde{\epsilon}_1|>\epsilon\sigma_n/\max_{1\le i\le n} i]}\} \\
&\quad \to 0\,.
\end{aligned}$$

Hence it follows that

$$n^{-1/2}(\overline{X}_n - (n+1)\mu/2) \to_d N(0, \sigma^2/3)\,.$$

Another nice way to view this is as follows: when $\beta = 1$, $X_i = \epsilon_1 + \ldots + \epsilon_i$, $i = 1, \ldots, n$. Since the X_i's have mean μ and variance σ^2,

$$X_i - i\mu = \tilde{\epsilon}_1 + \ldots + \tilde{\epsilon}_i \equiv S_i$$

where $\tilde{\epsilon}_i$ have mean 0 and variance σ^2. Thus

$$\begin{aligned}
n^{-1/2}(\overline{X}_n - (n+1)\mu/2) &= n^{-1/2}\left(n^{-1}\sum_{i=1}^n\sum_{j=1}^i \tilde{\epsilon}_j\right) = n^{-1}\sum_{i=1}^n n^{-1/2}S_i \\
&= \sigma\int_0^1 \mathbb{S}_n(t)dt
\end{aligned}$$

where

$$\mathbb{S}_n(t) \equiv \frac{1}{\sqrt{n}}\sum_{j=1}^{[nt]}\tilde{\epsilon}_j/\sigma$$

is the partial sum process of the $\tilde{\epsilon}_j/\sigma$'s. Now, as discussed in section 2.5, $\mathbb{S}_n \Rightarrow \mathbb{S}$ where $\mathbb{S}$ is a standard Brownian motion process; i.e. a mean zero Gaussian process with mean 0 and covariance function

$$Cov(\mathbb{S}(s), \mathbb{S}(t)) = s \wedge t\,.$$

Thus $\int_0^1 \mathbb{S}(t)dt$ is a linear combination of normal random variables, hence normal with

$$E\left\{\int_0^1 \mathbb{S}(t)dt\right\} = \int_0^1 E\{\mathbb{S}(t)\}dt = \int_0^1 0dt = 0\,,$$

and

$$\begin{aligned} Var(\int_0^1 \mathbb{S}(t)dt) &= E\left\{\int_0^1 \mathbb{S}(s)ds \int_0^1 \mathbb{S}(t)dt\right\} = E\left\{\int_0^1\int_0^1 \mathbb{S}(s)\mathbb{S}(t)dsdt\right\} \\ &= \int_0^1\int_0^1 E\{\mathbb{S}(s)\mathbb{S}(t)\}dsdt = \int_0^1\int_0^1 s\wedge t\, dsdt \\ &= 2\int_0^1 \left(\int_0^t sds\right)dt = \int_0^1 t^2 dt = 1/3\,. \end{aligned}$$

Thus

$$\sigma\int_0^1 \mathbb{S}_n(t)dt \to_d \sigma\int_0^1 \mathbb{S}(t)dt \sim N(0,\sigma^2/3)\,.$$

4. 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 $\sigma > 0$. Use (a) to find the limiting distribution of the sample *signal-noise ratio* $D_n \equiv \overline{X}_n/S_n$; i.e. show that $\sqrt{n}(D_n - d) \to_d N(0,V^2)$ with $d \equiv \mu/\sigma$ 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 $g(u,v) = u/\sqrt{v}$ is differentiable at points (u,v) with $v \neq 0$, and the derivative is $\nabla g(u,v) = (1/\sqrt{v}, u(-1/2)v^{-3/2})$ so that $\nabla g(\mu,\sigma^2) = (1/\sigma, (-1/2)\mu\sigma^{-3}) = (1/\sigma)(1, -(1/2)\mu/\sigma^2))$. Hence it follows from the delta method (g' theorem) that

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

and it is easy to calculate that

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

where $\gamma_1 \equiv \mu_3/\sigma^3$ and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$. Note that when the X_i's are normal (so $\gamma_1 = \gamma_2 = 0$), this reduces to $1 + d^2/2$. Thus under normality we have

$$\sqrt{n}(g(D_n) - g(d)) \to_d N(0,1)$$

if $g(x) \equiv \sqrt{2}\text{arcsinh}(x/\sqrt{2})$.

5. 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 the previous problem. We just need to proceed as in (b) of the previous problem 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

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

(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

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