**Solutions to Exercise Set 4.**

7.8. To find the asymptotic distribution of $\hat{\sigma}^2 = m_2 - (m_1 m_3/m_2)$, we need the asymptotic joint distribution of $(m_1, m_2, m_3)$. From the central limit theorem with $\mathrm{E}X = \mu_1 = 0$, we have

$$\sqrt{n}\left(\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} - \begin{pmatrix} 0 \\ \mu_2 \\ \mu_3 \end{pmatrix}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma\right)$$

where

$$\Sigma = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,X^2) & \mathrm{Cov}(X,X^3) \\ \mathrm{Cov}(X,X^2) & \mathrm{Var}(X^2) & \mathrm{Cov}(X^2,X^3) \\ \mathrm{Cov}(X,X^3) & \mathrm{Cov}(X^2,X^3) & \mathrm{Var}(X^3) \end{pmatrix}$$

Now apply Cramér's Theorem with $g(m_1, m_2, m_3) = m_2 - (m_1 m_3/m_2) = \hat{\sigma}^2$. We find $\dot{g}(m_1, m_2, m_3) = (-m_2/m_3, 1 + (m_1 m_3/m_2^2), -m_1/m_2)$ and $\dot{g}(0, \mu_2, \mu_3) = (-\mu_3/\mu_2, 1, 0)$. Using $\mathrm{Var}(X) = \mu_2$, $\mathrm{Cov}(X, X^2) = \mu_3$ and $\mathrm{Var}(X^2) = \mu_4 - \mu_2^2$, we find $\dot{g}\Sigma\dot{g}^{-1} = \mu_4 - \mu_2^2 - (\mu_3^2/\mu_2)$. Therefore, $\sqrt{n}(\hat{\sigma}^2 - \mu_2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2)$, where $\tau^2 = \mu_4 - \mu_2^2 - (\mu_3^2/\mu_2)$. This is less than or equal to $\mu_4 - \mu_2^2$ with equality if and only if $\mu_3 = 0$.

All two-point distributions with means zero are equivalent, up to change of scale, to one of the distributions, $\mathrm{P}(X = -1) = a/(a+1)$, $P(X = a) = 1/(a+1)$, for some $a > 0$. We find

$$\mathrm{E}X^2 = \frac{a^2}{a+1} + \frac{a}{a+1} = a$$
$$\mathrm{E}X^3 = \frac{a^3}{a+1} - \frac{a}{a+1} = a(a-1)$$
$$\mathrm{E}X^4 = \frac{a^4}{a+1} + \frac{a}{a+1} = a(a^2 - a + 1)$$

So $\tau^2 = \mu_4 - \mu_2^2 - (\mu_3^2/\mu_2) = a(a^2 - a + 1) - a^2 - a(a-1)^2 = 0$.

8.2. (a) Let us denote the product a×b central moment by $\mu_{ab} = \mathrm{E}[(X - \mu_x)^a (Y - \mu_y)^b]$, and its sample estimate by $m_{ab} = (1/n)\sum_1^n (X_i - \overline{X}_n)^a (Y_i - \overline{Y})^b$. From Theorem 8(a) we may deduce

$$\sqrt{n}\left[\begin{pmatrix} s_x^2 \\ s_y^2 \end{pmatrix} - \begin{pmatrix} \sigma_x^2 \\ \sigma_y^2 \end{pmatrix}\right] \xrightarrow{\mathcal{L}} \mathcal{N}(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C(XX,XX) & C(XX,YY) \\ C(XX,YY) & C(YY,YY) \end{pmatrix})$$

where $C(XX, XX) = \mu_{40} - \sigma_x^4$, $C(YY, YY) = \mu_{04} - \sigma_y^4$, and $C(XX, YY) = \mu_{22} - \sigma_x^2\sigma_y^2$. Now we use Cramér's Theorem with $g(x, y) = \log(x/y)$ and $\dot{g}(x, y) = (1/x, -1/y)$ to find

$$\sqrt{n}(Z_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \frac{\mu_{40}}{\sigma_x^4} - 2\frac{\mu_{22}}{\sigma_x^2\sigma_y^2} + \frac{\mu_{04}}{\sigma_y^4}).$$

(b) If we let $\gamma^2$ denote the variance of the limit of $\sqrt{n}(Z_n - \theta)$, then we may estimate $\gamma^2$ by $\hat{\gamma}^2 = b_{40} - 2b_{22} + b_{04}$, where $b_{ab} = m_{ab}/(s_x^a s_y^b)$. Then an asymptotically distribution-free confidence interval for $\theta$ at level $1 - \alpha$ is $Z_n - z_{\alpha/2}\hat{\gamma}/\sqrt{n} < \theta < Z_n + z_{\alpha/2}\hat{\gamma}/\sqrt{n}$, where $z_{\alpha/2}$ is the $\alpha/2$ cutoff point for the standard normal distribution.

9.6. $X$ and $Y$ are independent if and only if $\mathrm{P}(X=1, Y=1) = \mathrm{P}(X=1)\mathrm{P}(Y=1)$. This becomes $p_{11} = (p_{11}+p_{12})(p_{11}+p_{21})$, or $p_{11} = p_{11}^2 + p_{11}p_{21} + p_{11}p_{12} + p_{12}p_{21}$, or $p_{12}p_{21} = p_{11}(1-p_{11}-p_{12}-p_{21}) = p_{11}p_{22}$. This is the equation $\theta = 1$.

(a) The asymptotic distribution of $\hat{\boldsymbol{p}} = (\hat{p}_{11}, \hat{p}_{12}, \hat{p}_{21}, \hat{p}_{22})$ is given by

$$\sqrt{n}(\hat{\boldsymbol{p}} - \boldsymbol{p}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{P} - \boldsymbol{p}\boldsymbol{p}^T)$$

so, using $g(\boldsymbol{p}) = p_{11}p_{22}/(p_{12}p_{21})$, we have

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \dot{g}(\boldsymbol{p})^T(\boldsymbol{P} - \boldsymbol{p}\boldsymbol{p}^T)\dot{g}(\boldsymbol{p}))$$

One finds that

$$\dot{g}(\boldsymbol{p})^T = \theta\left(\frac{1}{p_{11}}, \frac{-1}{p_{12}}, \frac{-1}{p_{21}}, \frac{1}{p_{22}}\right).$$

Therefore, $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$, where

$$\sigma^2 = \dot{g}(\boldsymbol{p})^T(\boldsymbol{P} - \boldsymbol{p}\boldsymbol{p}^T)\dot{g}(\boldsymbol{p}) = \dot{g}(\boldsymbol{p})^T\boldsymbol{P}\dot{g}(\boldsymbol{p}) = \theta^2\left(\frac{1}{p_{11}} + \frac{1}{p_{12}} + \frac{1}{p_{21}} + \frac{1}{p_{22}}\right)$$

(b) Now let $g(\theta) = \log(\theta)$, with $\dot{g}(\theta) = 1/\theta$. We have

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2/\theta^2) = \mathcal{N}(0, \left(\frac{1}{p_{11}} + \frac{1}{p_{12}} + \frac{1}{p_{21}} + \frac{1}{p_{22}}\right)).$$