# Statistics 581, Problem Set 3 Solutions 

Wellner; 10/18/2018

1. Ferguson, ACILST, page 34, problem 1(a) (modified slightly):

Suppose that $X_{1}, X_{2}, \ldots$ are i.i.d. in $R^{2}$ with distribution giving probability $\theta_{1}$ to $(1,0)^{\prime}$, probability $\theta_{2}$ to $(0,1)^{\prime}, \theta_{3}$ to $(-1,0)^{\prime}$ and $\theta_{4}$ to $(0,-1)^{\prime}$ where $\theta_{j} \geq 0$ for $j=1,2,3,4$ and $\theta_{1}+\theta_{2}+\theta_{3}+\theta_{4}=1$.
(a) Find $\mu=E\left(X_{1}\right)$.
(b) Compute $E\left(X_{1} X_{1}^{T}\right)$ and $\Sigma=E\left(X_{1}-\mu\right)\left(X_{1}-\mu\right)^{T}$.
(c) Find the limiting distribution of $\sqrt{n}\left(\bar{X}_{n}-\mu\right)$ and describe the resulting approximation to the distribution of $\bar{X}_{n}$.
(d) Find values of $\left(\theta_{1}, \ldots, \theta_{4}\right)$ such that $\Sigma$ has rank 1 and $\operatorname{det}(\Sigma)=0$.

Solution: (a) The mean vector $\mu$ is

$$
\mu=E\left(X_{1}\right)=\theta_{1}\binom{1}{0}+\theta_{2}\binom{0}{1}+\theta_{3}\binom{-1}{0}+\theta_{4}\binom{0}{-1}=\binom{\theta_{1}-\theta_{3}}{\theta_{2}-\theta_{4}} .
$$

(b) The second moment matrix $E\left(X X^{T}\right)$ is

$$
\begin{aligned}
E\left(X X^{T}\right) & =\theta_{1}\left(\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right)+\theta_{2}\left(\begin{array}{ll}
0 & 0 \\
0 & 1
\end{array}\right)+\theta_{3}\left(\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right)+\theta_{4}\left(\begin{array}{ll}
0 & 0 \\
0 & 1
\end{array}\right) \\
& =\left(\begin{array}{cc}
\theta_{1}+\theta_{3} & 0 \\
0 & \theta_{2}+\theta_{4}
\end{array}\right)
\end{aligned}
$$

and hence the covariance matrix $\Sigma=E\left(X X^{T}\right)-E(X) E(X)^{T}$ is

$$
\begin{aligned}
\Sigma & =\left(\begin{array}{cc}
\theta_{1}+\theta_{3} & 0 \\
0 & \theta_{2}+\theta_{4}
\end{array}\right)-\left(\begin{array}{cc}
\left(\theta_{1}-\theta_{3}\right)^{2} & \left(\theta_{1}-\theta_{3}\right)\left(\theta_{2}-\theta_{4}\right. \\
\left(\theta_{1}-\theta_{3}\right)\left(\theta_{2}-\theta_{4}\right) & \left(\theta_{2}-\theta_{4}\right)^{2}
\end{array}\right) \\
& =\left(\begin{array}{cc}
\theta_{1}+\theta_{3}-\left(\theta_{1}-\theta_{3}\right)^{2} & -\left(\theta_{1}-\theta_{3}\right)\left(\theta_{2}-\theta_{4}\right) \\
-\left(\theta_{1}-\theta_{3}\right)\left(\theta_{2}-\theta_{4}\right) & \left(\theta_{2}+\theta_{4}\right)-\left(\theta_{2}-\theta_{4}\right)^{2}
\end{array}\right)
\end{aligned}
$$

(c) By the multivariate CLT, the limiting distribution of $\sqrt{n}\left(\bar{X}_{n}-\mu\right)$ is $N_{2}(0 \Sigma)$. Thus the approximating normal distribution of $\bar{X}_{n}$ is centered at $\mu$ with covariance matrix $n^{-1} \Sigma$.
(d) The determinant of $\Sigma$ is given by
$\operatorname{det}(\Sigma)=\left\{\left(\theta_{1}+\theta_{3}\right)-\left(\theta_{1}-\theta_{3}\right)^{2}\right\}\left\{\left(\theta_{2}+\theta_{4}\right)-\left(\theta_{2}-\theta_{4}\right)^{2}\right\}-\left(\theta_{1}-\theta_{3}\right)^{2}\left(\theta_{2}-\theta_{4}\right)^{2}$.
This equals 0 if $\theta_{2}=\theta_{4}=0$ or if $\theta_{1}=\theta_{3}=0$. In the first case

$$
\Sigma=\left(\begin{array}{cc}
\theta_{1}+\theta_{3}-\left(\theta_{1}-\theta_{3}\right)^{2} & 0 \\
0 & 0
\end{array}\right)=\left(\begin{array}{cc}
1-\left(\theta_{1}-\theta_{3}\right)^{2} & 0 \\
0 & 0
\end{array}\right)
$$

and in the second case

$$
\Sigma=\left(\begin{array}{cc}
0 & 0 \\
0 & \left(\theta_{2}+\theta_{4}\right)-\left(\theta_{2}-\theta_{4}\right)^{2}
\end{array}\right)=\left(\begin{array}{cc}
0 & 0 \\
0 & 1-\left(\theta_{2}-\theta_{4}\right)^{2}
\end{array}\right) .
$$

2. (Van der Vaart, page 24)
(a) Suppose that $X_{n}$ and $Y_{n}$ are independent random vectors with $X_{n} \rightarrow_{d} X$ and $Y_{n} \rightarrow_{d} Y$. Show that $\left(X_{n}, Y_{n}\right) \rightarrow_{d}(X, Y)$ where $X$ and $Y$ are independent.
(b) Suppose that $P\left(X_{n}=i / n\right)=1 / n$ for $i=1,2, \ldots, n$. show that $X_{n} \rightarrow_{d} X \sim$ Uniform $(0,1)$.
(c) Consider the $X_{n}$ 's as in (b). Show that there exists a Borel set $B$ such that $P\left(X_{n} \in B\right)=1$ but $P(X \in B)=0$. In particular, with $P_{n}=\mathcal{L}\left(X_{n}\right)$ and $P=\mathcal{L}(X)$, $d_{T V}\left(P_{n}, P\right)=1$ for each $n$.

Solution:(a) Since $X_{n} \rightarrow_{d} X$ in $\mathbb{R}^{m}$ and $Y_{n} \rightarrow_{d} Y$ in $\mathbb{R}^{k}$ (as random vectors), we know that their characteristic functions also converge pointwise: for every $s \in \mathbb{R}^{m}$, $t \in \mathbb{R}^{k}$,

$$
\begin{aligned}
\phi_{X_{n}}(s)=E e^{i s^{T} X_{n}} & \rightarrow E e^{i s^{T} X}=\phi_{X}(s) \text { and } \\
\phi_{Y_{n}}(t)=E e^{i t^{T} Y_{n}} & \rightarrow E e^{i t^{T} Y}=\phi_{Y}(t)
\end{aligned}
$$

Then, since $X_{n}$ and $Y_{n}$ are independent,

$$
\begin{aligned}
\phi_{X_{n}, Y_{n}}(s, t) & =E e^{i s^{T} X_{n}+t^{T} Y_{n}}=E e^{i s^{T} X_{n}} E e^{i t^{T} Y_{n}} \\
& \rightarrow \phi_{X}(s) \cdot \phi_{Y}(t) .
\end{aligned}
$$

This implies that $\left(X_{n}, Y_{n}\right) \rightarrow_{d}(X, Y)$ where $X$ and $Y$ are independent.
(b) Note that for each $x \in[0,1]$

$$
F_{n}(x)=P\left(X_{n} \leq x\right)=\lfloor n x\rfloor / n \rightarrow x=P(X \leq x)
$$

It follows that $X_{n} \rightarrow X \sim \operatorname{Uniform}(0,1)$.
(c) Let $B=\{1 / n, \ldots, n / n=1\}$. Then $P\left(X_{n} \in B\right)=1$ but $P(X \in B)=0$. Thus $d_{T V}\left(P_{n}, P\right)=1$.
3. ACILST, page 34 , problem 5: Suppose that $X_{1}, X_{2}, \ldots$ are i.i.d. random variables with mean $\mu$ and variance $\sigma^{2}<\infty$. Let $T_{n}=\sum_{j=1}^{n} z_{n j} X_{j}$ where the $\left\{z_{n j}\right\}_{j=1}^{n}$ are given numbers. Let $\mu_{n}=E\left(T_{n}\right)$ and $\sigma_{n}^{2}=\operatorname{Var}\left(T_{n}\right)$. Use the LindebergFeller central limit theorem to show that $\left(T_{n}-\mu_{n}\right) / \sigma_{n} \rightarrow_{d} Z \sim N(0,1)$ if $\max _{1 \leq j \leq n} z_{n j}^{2} / \sum_{j=1}^{n} z_{n j}^{2} \rightarrow 0$ as $n \rightarrow \infty$.

Solution: Let $Y_{i} \equiv X_{i}-\mu$ so that $E\left(Y_{i}\right)=0$ and $\operatorname{Var}\left(Y_{i}\right)=\sigma^{2}$. We also let $X_{n} i \equiv z_{n i} Y_{i}$ so that $T_{n}-\mu_{n}=\sum_{i=1}^{n} X_{n i}$, and note that $\sigma_{n}^{2}=\operatorname{Var}\left(T_{n}\right)=\sigma^{2} \sum_{i=1}^{n} z_{n i}^{2}$. To prove the claimed asymptotic normality we need to check that the Lindebergcondition holds. For $\epsilon>0$ we want to show that

$$
L F_{n}(\epsilon)=\frac{1}{\sigma_{n}^{2}} \sum_{i=1}^{n} E\left\{X_{n i}^{2} 1\left\{\left|X_{n i}\right|>\epsilon \sigma_{n}\right\} \rightarrow 0\right.
$$

for every $\epsilon>0$. But we note that

$$
L F_{n}(\epsilon)=\frac{1}{\sigma_{n}^{2}} \sum_{i=1}^{n} E\left\{z_{n i}^{2} Y_{i}^{2} 1\left\{\left|z_{n i}\right|\left|Y_{i}\right|>\epsilon \sigma\left(\sum_{i=1}^{n} z_{n i}^{2}\right)^{1 / 2}\right\}\right\}
$$

$$
\begin{aligned}
& \leq \frac{1}{\sigma_{n}^{2}} \sum_{i=1}^{n} z_{n i}^{2} E\left\{Y_{i}^{2} 1\left\{\max _{1 \leq i \leq n}\left|z_{n i}\right|\left|Y_{i}\right|>\epsilon \sigma\left(\sum_{i=1}^{n} z_{n i}^{2}\right)^{1 / 2}\right\}\right\} \\
& =\frac{1}{\sigma^{2}} E\left\{Y_{1}^{2} 1\left\{\left|Y_{1}\right|^{2}>\epsilon \sigma \sqrt{\left.\sum_{i=1}^{n} z_{n i}^{2} / \max _{1 \leq i \leq n}\left|z_{n i}\right|\right\}}\right\}\right. \\
& \rightarrow 0
\end{aligned}
$$

by the dominated convergence theorem with dominating function $Y_{1}^{2}$ since $E\left(Y_{1}^{2}\right)<$ $\infty$.
4. (a) ACILST, problem 4, page 49: Let $X_{1}, \ldots, X_{n}$ be a sample of size $n$ from the beta distribution Beta $(\theta, 1)$ with $\theta>0$. Show that the method of moments estimate of $\theta$ is $\hat{\theta}_{n}=\bar{X}_{n} /\left(1-\bar{X}_{n}\right)$.
(b) Find the asymptotic distribution of $\hat{\theta}_{n}$.
(c) Is $\hat{\theta}_{n}$ asymptotically linear? If so, find the influence function of $\hat{\theta}_{n}$.
(d) Find the Cramér-Rao lower bound for estimation of $\theta$ and compare it to the asymptotic variance you found in (b).

Solution: (a) If $Y \sim \operatorname{Beta}(\alpha, \beta)$, then $E(Y)=\alpha /(\alpha+\beta)$, so $E_{\theta} X_{1}=\theta /(\theta+1)$. Thus the method of moments estimator $\hat{\theta}_{n}$ of $\theta$ satisfies $\bar{X}_{n}=\hat{\theta}_{n} /\left(\hat{\theta}_{n}+1\right)$. Thus $\hat{\theta}_{n}=\bar{X}_{n} /\left(1-\bar{X}_{n}\right)$ as claimed.
(b) Now $\operatorname{Var}(Y)=\alpha \beta /\left((\alpha+\beta)^{2}(\alpha+\beta+1)\right)$, so $\operatorname{Var}_{\theta}\left(X_{1}\right)=\theta /\left((\theta+1)^{2}(\theta+2)\right)$. Thus $\sqrt{n}\left(\bar{X}_{n}-E_{\theta}\left(X_{1}\right)\right) \rightarrow_{d} N\left(0, \operatorname{Var}_{\theta}\left(X_{1}\right)\right)$, and by the delta method with $g(y) \equiv$ $y /(1-y)$ we find that
$\sqrt{n}\left(\hat{\theta}_{n}-\theta\right) \rightarrow_{d} \dot{g}\left(E_{\theta}\left(X_{1}\right)\right) N\left(0, \frac{\theta}{(\theta+1)^{2}(\theta+2)}\right)=N\left(0, \frac{\theta(\theta+1)^{2}}{(\theta+2)}\right) \equiv N\left(0, V_{\theta}^{2}\right)$.
(c) Note that the density $f_{\theta}$ of one observation is $f_{\theta}(x)=\theta x^{\theta-1} 1_{(0,1)}(x)$. Thus $\log f_{\theta}(x)=\log \theta+(\theta-1) \log x$, and the score function for $\theta$ is $\mathrm{i}_{\theta}(x)=\theta^{-1}+\log x$. This yields $I_{\theta}=E\left\{-\ddot{\mathrm{l}}_{\theta \theta}\left(X_{1}\right)\right\}=\theta^{-2}$. Thus the Cramér - Rao bound for unbiased estimators $\tilde{\theta}_{n}$ of $\theta$ is given by $\operatorname{Var}\left(\sqrt{n}\left(\tilde{\theta}_{n}-\theta\right)\right) \geq \theta^{2} \equiv C R L B(\theta)$. Note that this is strictly smaller than the asymptotic variance of the method of moments estimator $\hat{\theta}_{n}$ found in (b). In fact $\hat{\theta}_{n}$ is severely inefficient for small values of $\theta$ since the ratio $C R L B(\theta) / V_{\theta}^{2} \rightarrow 0$ as $\theta \rightarrow 0$; see Figure 1.
(d) From the extended delta method discussed in class on 10 October, with $g(v)=$ $v /(1-v), g^{\prime}(v)=1 /(1-v)^{2}$ it follows that

$$
\sqrt{n}\left(g\left(\bar{X}_{n}\right)-g\left(E_{\theta}(X)\right)-g^{\prime}\left(E_{\theta}(X)\right)\left(\bar{X}_{n}-E_{\theta}(X)\right)=o_{p}(1) .\right.
$$

Since $g\left(\bar{X}_{n}\right)=\hat{\theta}_{n}$ and $g\left(E_{\theta}(X)\right)=\theta$ it follows that

$$
\begin{aligned}
\sqrt{n}\left(\hat{\theta}_{n}-\theta\right) & =\frac{1}{\left(1-E_{\theta}(X)\right)^{2}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n}\left(X_{i}-E_{\theta}(X)\right)+o_{p}(1) \\
& =\frac{1}{\sqrt{n}} \sum_{i=1}^{n}(1+\theta)^{2}\left(X_{i}-E_{\theta}(X)\right)+o_{p}(1) .
\end{aligned}
$$

Thus $\hat{\theta}_{n}$ is asymptotically linear with influence function

$$
\psi(x) \equiv \psi_{\theta}(x)=(1+\theta)^{2}\left(x-E_{\theta}(X)\right)
$$



Figure 1: Ratio $C R L B(\theta) / V_{\theta}^{2}$ as a function of $\theta$
5. Suppose that $X_{1}, \ldots, X_{n}$ are i.i.d. with $E\left(X_{1}\right)=\mu, \operatorname{Var}\left(X_{1}\right)=\sigma^{2}<\infty$, and $E\left|X_{1}\right|^{6}<\infty$. Let $M_{j, n} \equiv n^{-1} \sum_{i=1}^{n}\left(X_{i}-\bar{X}_{n}\right)^{j}$ be the $j$-th sample central moment for $j \in\{2,3, \ldots\}$.
(a) ACILST, page 49, problem 3.
(b) Find the joint asymptotic distribution of $\sqrt{n}\left(\left(\bar{X}_{n}, M_{2, n}, M_{3, n}\right)^{T}-\left(\mu, \sigma^{2}, m_{3}\right)^{T}\right)$ where $m_{3} \equiv E\left(X_{1}-\mu\right)^{3}$ is the population 3rd central moment.
(c) Find the asymptotic distribution of $\sqrt{n}\left(\kappa_{3, n}-\kappa_{3}\right)$ where $\kappa_{3, n} \equiv M_{3, n} / M_{2, n}^{3 / 2}$ is the sample skewness and $\kappa_{3} \equiv m_{3} / \sigma^{3}$ is the population skewness. (See vdV Example 3.5, page 29.)

Solution: (a) $M_{3, n}$ is location invariant, so we may assume without loss of generality that $\mu=E\left(X_{1}\right)=0$. As a first step, note that

$$
\begin{aligned}
M_{3, n} & =n^{-1} \sum_{i=1}^{n}\left(X_{i}^{3}-3 X_{i}^{2} \cdot \bar{X}_{n}+3 X_{i} \bar{X}_{n}-\bar{X}_{n}^{3}\right) \\
& =\bar{X}^{3}{ }_{n}-3 \bar{X}^{2}{ }_{n}+2 \bar{X}_{n}^{3} \\
& \rightarrow_{p} E\left(X^{3}\right)-3 E\left(X^{2}\right) \cdot E(X)+2(E(X))^{3} \\
& =E\left(X^{3}\right)-0+0=E\left(X^{3}\right)
\end{aligned}
$$

by the WLLN and the Mann-Wald theorem. Now by the multivariate CLT, assuming that $E\left(X^{6}\right)<\infty$

$$
\sqrt{n}\left(\begin{array}{c}
\bar{X}_{n}-E(X) \\
\bar{X}^{2} \\
\\
\bar{X}_{n}^{3}-E\left(X^{2}\right) \\
\left.\bar{X}^{3}\right)
\end{array}\right) \rightarrow_{d} \underline{Z} \sim N_{3}(0, \Sigma)
$$

where, with $\mu_{k} \equiv E\left(X^{k}\right)$,

$$
\Sigma=\left(\begin{array}{ccc}
\mu_{2} & \mu_{3} & \mu_{4} \\
\mu_{3} & \mu_{4}-\mu_{2}^{2} & \mu_{5}-\mu_{2} \mu_{3} \\
\mu_{4} & \mu_{5}-\mu_{2} \mu_{3} & \mu_{6}-\mu_{3}^{2}
\end{array}\right)
$$

Furthermore,

$$
\sqrt{n}\left(M_{3, n}-m_{3}\right)=\sqrt{n}\left(g\left(\bar{X}_{n},{\overline{X^{2}}}_{n},{\overline{X^{3}}}_{n}\right)-g\left(0, \mu_{2}, \mu_{3}\right)\right)
$$

where

$$
g(u, v, w)=w-3 v u+2 u^{3}
$$

is differentiable at $(u, v, w)=\left(0, \mu_{2}, \mu_{3}\right)$ with

$$
g^{\prime}(u, v, w)=\left(-3 v+6 u^{2},-3 u, 1\right), \text { so that } g^{\prime}\left(0, \mu_{2}, \mu_{3}\right)=\left(-3 \mu_{2}, 0,1\right) .
$$

It follows from the delta-method (or $g$-prime theorem) that, with $\sigma^{2} \equiv \mu_{2}$,

$$
\begin{aligned}
\sqrt{n}\left(M_{3, n}-m_{3}\right) & =\sqrt{n}\left(g\left(\bar{X}_{n},{\overline{X^{2}}}_{n},{\overline{X^{3}}}_{n}\right)-g\left(0, \mu_{2}, \mu_{3}\right)\right) \\
& \rightarrow_{d}\left(-3 \mu_{2}, 0,1\right) \underline{Z}=Z_{3}-3 \mu_{2} Z_{1} \\
& \sim N\left(0, \mu_{6}+9 \sigma^{6}-6 \sigma^{2} \mu_{4}\right) .
\end{aligned}
$$

Note that when $X \sim N\left(0, \sigma^{2}\right)$ with $\mu_{4}=3 \sigma^{4}$ this becomes $N\left(0,6 \sigma^{6}\right)$.
(b) Now we regard $g$ in (a) above as $g_{3}$ where $g=\left(g_{1}, g_{2}, g_{3}\right)^{T}$ with

$$
g_{1}(u, v, w)=u, \quad g_{2}(u, v, w)=v-u^{2}, \quad g_{3}(u, v, w)=w-3 v u+2 u^{3} .
$$

Thus

$$
\begin{aligned}
\nabla g_{1}(u, v, w) & =(1,0,0)^{T} \\
\nabla g_{2}(u, v, w) & =(-2 u, 1,0)^{T} \\
\nabla g_{3}(u, v, w) & =\left(-3 v+6 u^{3},-3 u, 1\right)^{T}
\end{aligned}
$$

and we find that

$$
g^{\prime}\left(0, \mu_{2}, \mu_{3}\right)=\left(\begin{array}{ccc}
1 & 0 & 0 \\
0 & 1 & 0 \\
-3 \mu_{2} & 0 & 1
\end{array}\right)
$$

Thus by the delta-method

$$
\sqrt{n}\left(\begin{array}{c}
\bar{X}_{n}-0 \\
M_{2, n}-\sigma^{2} \\
M_{3, n}-m_{3}
\end{array}\right) \rightarrow_{d}\left(\begin{array}{c}
Z_{1} \\
Z_{2} \\
Z_{3}-3 \sigma^{2} Z_{1}
\end{array}\right) \equiv \underline{\widetilde{Z}} \sim N_{3}(0, \widetilde{\Sigma})
$$

where

$$
\begin{aligned}
\widetilde{\Sigma} & =\left(\begin{array}{ccc}
\sigma^{2} & \mu_{3} & \mu_{4}-3 \sigma^{4} \\
\mu_{3} & \mu_{4}-\sigma^{4} & \mu_{5}-3 \sigma^{2} \mu_{3} \\
\mu_{4}-3 \sigma^{4} & \mu_{5}-3 \sigma^{2} \mu_{3} & \mu_{6}+9 \sigma^{6}-6 \sigma^{2} \mu_{4}
\end{array}\right) \\
& =\left(\begin{array}{ccc}
\sigma^{2} & \mu_{3} & \sigma^{4} \gamma_{2} \\
\mu_{3} & \sigma^{4}\left(2+\gamma_{2}\right) & \sigma^{5}\left(\mu_{5} / \sigma^{5}-3 \kappa_{3}\right) \\
\sigma^{4} \gamma_{2} & \sigma^{5}\left(\mu_{5} / \sigma^{5}-3 \kappa_{3}\right) & \sigma^{6}\left(\mu_{6} / \sigma^{6}-9-6 \gamma_{2}\right.
\end{array}\right) .
\end{aligned}
$$

Note that $\mu_{4}-3 \sigma^{4}=\sigma^{4}\left(\mu_{4} / \sigma^{4}-3\right)=\sigma^{4} \gamma_{2}$ in the notation of the 581 course notes, example 2.3.2 (so that $\mu_{3}-\sigma^{4}=\sigma^{4}\left(\mu_{3} / \sigma^{4}-1\right)=\sigma^{4}\left(2+\gamma_{2}\right)$ ), while

$$
\begin{aligned}
\mu_{5}-3 \sigma^{2} \mu_{3} & =\sigma^{5}\left(\mu_{5} / \sigma^{5}-3 \mu_{3} / \sigma^{3}\right) \\
\mu_{6}+9 \sigma^{6}-6 \sigma^{2} \mu_{4} & =\sigma^{6}\left(\mu_{6} / \sigma^{6}+9-6 \mu_{4} / \sigma^{4}\right)
\end{aligned}
$$

(c) To find the limiting distribution of $\kappa_{3, n}$ we use the delta method again in combination with the result of (b): now $g(u, v, w)=w / v^{3 / 2}$ is differentiable at $v \neq 0$ with derivative $g^{\prime}(u, v, w)=\left(0,(-3 / 2) v^{-5 / 2} w, v^{-3 / 2}\right)$ so that

$$
g^{\prime}\left(0, \sigma^{2}, m_{3}\right)=\left(0,-(3 / 2) \sigma^{-2}\left(m_{3} / \sigma^{3}\right), \sigma^{-3}\right)=\left(0,-(3 / 2) \sigma^{-2} \kappa_{3}, \sigma^{-3}\right)
$$

Thus it follows from the delta method that

$$
\begin{aligned}
\sqrt{n}\left(\kappa_{3, n}-\kappa_{3}\right) & =\sqrt{n}\left(g\left(\bar{X}_{n}, M_{2, n}, M_{3, n}\right)-g\left(0, \sigma^{2}, m_{3}\right)\right) \\
& \rightarrow_{d} \sigma^{-3} \widetilde{Z}_{3}-(3 / 2) \sigma^{-2} \kappa_{3} \widetilde{Z}_{2} \\
& \sim N\left(0, V^{2}\right)
\end{aligned}
$$

where

$$
V^{2}=\frac{\mu_{6}}{\sigma^{6}}-9-6 \gamma_{2}+\frac{9}{4} \kappa_{3}^{2}\left(2+\gamma_{2}\right)-3 \kappa_{3}\left(\frac{\mu_{5}}{\sigma^{5}}-3 \kappa_{3}\right) .
$$

Note that when $X \sim N\left(0, \sigma^{2}\right)$ we have $\kappa_{3}=0, \gamma_{2}=0$ and $\mu_{6} / \sigma^{6}=15$, so that $V^{2}=15-9=6$. Thus if $X \sim N\left(\mu, \sigma^{2}\right)$ the test "reject $H: X \sim N\left(\mu, \sigma^{2}\right)$ in favor of $K_{\text {skew }}: \quad \kappa_{3} \neq 0$ if $\left|\sqrt{n} \kappa_{3, n}\right|>\sqrt{6} z_{\alpha / 2}$ " has approximate size $\alpha$ for large $n$; i.e. $P_{\text {norm }}\left(\left|\sqrt{n} \kappa_{3, n}\right|>\sqrt{6} z_{\alpha / 2}\right) \rightarrow P\left(|\sqrt{6} Z|>\sqrt{6} z_{\alpha / 2}\right)=\alpha$ as $n \rightarrow \infty$.
The following development was not part of the problem as stated, but I discussed it in class on 17 October. If the distribution function $F$ of $X$ has $\kappa_{3} \neq 0$, then (assuming that $E|X|^{6}<\infty$ )

$$
P_{F}\left(\left|\sqrt{n} \kappa_{3, n}\right|>\sqrt{6} z_{\alpha / 2}\right)=P_{F}\left(\left|\sqrt{n}\left(\kappa_{3, n}-\kappa_{3}\right)+\sqrt{n} \kappa_{3}\right|>\sqrt{6} z_{\alpha / 2}\right) \rightarrow 1
$$

as $n \rightarrow \infty$ since $\left|\sqrt{n} \kappa_{3}\right| \rightarrow \infty$ and $\left|\sqrt{n}\left(\kappa_{3, n}-\kappa_{3}\right)\right|=O_{p}(1)$. What about the (local asymptotic power) of this test? For example, what if $X \sim F_{\alpha}$ with density $f(x ; \alpha)=2 \phi(x) \Phi(\alpha x)$ ? This $f$ is the skew-normal family of densities. Figure 2 shows the densities $f(x ; \alpha)$ for $\alpha \in\{0,1,3,10\}$ (in blue, magenta, purple, and green).

It turns out that the skewness of this family is given by

$$
\kappa_{3}\left(F_{\alpha}\right)=\sqrt{2}(4-\pi) \alpha^{3} /\left(\pi+(\pi-2) \alpha^{2}\right)^{3 / 2}
$$

see Azzalini (2014), The Skew-Normal and Related Families, pages 30-31. Figure 2 shows the densities $f(x ; \alpha)$ for $\alpha \in\{0,1,3,10\}$ (in blue, magenta, purple, and green). This function of $\alpha$ has first two derivatives equal to zero at $\alpha=0$, and it seems natural to reparametrize by $\alpha(\beta)=\beta^{1 / 3}$. See Figures 3 and 4. Then with

$$
\kappa_{3}\left(\tilde{F}_{\beta}\right)=\kappa_{3}\left(F_{\beta^{1 / 3}}\right)=\frac{\sqrt{2}(4-\pi) \beta}{\left(\pi+(\pi-2) \beta^{2 / 3}\right)^{3 / 2}}
$$



Figure 2: The densities $f(x ; \alpha)$ for $\alpha \in\{0,1,3,10\}$
it follows that

$$
\sqrt{n} \kappa_{3}\left(\tilde{F}_{b / \sqrt{n}}\right)=\frac{\sqrt{2}(4-\pi) b}{\left(\pi+(\pi-2)\left(b / n^{1 / 2}\right)^{2 / 3}\right)^{3 / 2}} \rightarrow \frac{\sqrt{2}(4-\pi) b}{\pi^{3 / 2}} .
$$

Then, modulo an argument using a triangular array multivariate CLT or reasoning via contiguity theory (see Chapter 3),

$$
\begin{aligned}
& P_{\tilde{F}_{b / \sqrt{n}}}\left(\left|\sqrt{n} \kappa_{3, n}\right|>\sqrt{6} z_{\alpha / 2}\right) \\
& \quad=P_{\tilde{F}_{b / \sqrt{n}}}\left(\left|\sqrt{n}\left(\kappa_{3, n}-\kappa_{3}\left(\tilde{F}_{b / n^{1 / 2}}\right)\right)+\sqrt{n} \kappa_{3}\left(\tilde{F}_{b / n^{1 / 2}}\right)\right|>\sqrt{6} z_{\alpha / 2}\right) \\
& \quad \rightarrow P\left(\left|\sqrt{6} Z+\sqrt{2}(4-\pi) b /(\pi-2)^{3 / 2}\right|>\sqrt{6} z_{\alpha / 2}\right) .
\end{aligned}
$$

Thus we can approximate the power of the skewness test of normality for these particular skew-normal (local) alternatives.


Figure 3: The function $\alpha \mapsto \kappa_{3}\left(F_{\alpha}\right)$


Figure 4: The function $\beta \mapsto \kappa_{3}\left(\tilde{F}_{\beta}\right)$

