MA 413 - Statistics for Data Science

Solutions to Exercise 7

1. We write first $Y_i = X_i X_{i+1}$ for $i = 1, 3, 5, \dots, 2n-1$. Then Y_i are iid with mean $\mathbb{E}[Y_i] = \mathbb{E}[X_i]\mathbb{E}[X_{i+1}] = 0$ and variance $\mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2 = 1$. By CLT we have
$$\frac{\sum_{i=1}^n Y_i}{\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$$
and therefore $\left(\frac{\sum_{i=1}^n Y_i}{\sqrt{n}}\right)^2 \xrightarrow{d} \chi_1^2$. Further, we have that $X_i^2 \sim \chi_1^2$ and by law of large numbers
$$\frac{1}{2n}\sum_{i=1}^{2n} X_i^2 \xrightarrow{p} 1$$
and so $\frac{1}{n}\sum_{i=1}^{2n} X_i^2 \xrightarrow{p} 2$ and $\left(\frac{1}{n}\sum_{i=1}^{2n} X_i^2\right)^2 \xrightarrow{p} 4$. Finally from the general version of Slutsky's theorem we have that
$$\frac{n\left(\sum_{i=1}^n Y_i\right)^2}{\left(\sum_{i=1}^{2n} X_i^2\right)^2} \xrightarrow{d} 4\chi_1^2$$

2. For discrete random variables it suffices to show convergence of the probability mass functions. We have
$$\begin{aligned} P(X_n = t) &= \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} = \frac{n!}{t!(n-t)!}\frac{1}{(n-1)^t}\left(1-\frac{1}{n}\right)^n \\ &= \frac{1}{t!}\frac{n\cdot\dots\cdot(n-t+1)}{(n-1)^t}\left(1-\frac{1}{n}\right)^n \xrightarrow{n\to\infty} \frac{1}{t!}e^{-1}, \end{aligned}$$
which is the pmf of a Poisson(1) distribution.

3. We are working with moment generating functions. The MGF of a Beta($\frac{1}{n}, \frac{1}{n}$) is
$$\begin{aligned} M_{Y_n}(t) &= 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\frac{1}{n}+r}{\frac{2}{n}+r}\right)\frac{t^k}{k!} \\ &= 1 + \sum_{k=1}^{\infty}\left(\frac{1}{2}\prod_{r=1}^{k-1}\frac{\frac{1}{n}+r}{\frac{2}{n}+r}\right)\frac{t^k}{k!} \end{aligned}$$
and for $n \to \infty$ we get
$$M_{Y_n}(t) \to 1 + \frac{1}{2}\sum_{k=1}^{\infty}\frac{t^k}{k!} = 1 + \frac{1}{2}\left(e^t - 1\right) = \frac{1}{2} + \frac{1}{2}e^t$$
that is the MGF of a Bernoulli(1/2) r.v.

4. Let Y be the random variable corresponding to the number of U_i's that are below x. Then

$$F_{U_{(m+1)}}(x) = P(Y \geq m+1) = \sum_{k=m+1}^{2m+1} \binom{2m+1}{k} F(x)^k (1-F(x))^k$$

and the density function is

$$\begin{aligned}
f_{U_{(m+1)}}(x) &= F'_{U_{(m+1)}}(x) \\
&= \sum_{k=m+1}^{2m+1} \binom{2m+1}{k} kF(x)^{k-1} (1-F(x))^{2m+1-k} f(x) \\
&- \sum_{k=m+1}^{2m+1} \binom{2m+1}{k} (2m+1-k)F(x)^k (1-F(x))^{2m+1-k-1} f(x) \\
&= \binom{2m+1}{m+1} (m+1) f(x) F(x)^m (1-F(x))^m \\
&+ \sum_{k=m+2}^{2m+1} \binom{2m+1}{k} kF(x)^{k-1} (1-F(x))^{2m+1-k} f(x) \\
&- \sum_{k=m+1}^{2m} \binom{2m+1}{k} (2m+1-k)F(x)^k (1-F(x))^{2m+1-k-1} f(x) \\
&= \binom{2m+1}{m+1} (m+1) f(x) F(x)^m (1-F(x))^m .
\end{aligned}$$

Taking $F(x) = x$ and $f(x) = 1$ for uniform$(0,1)$ we get

$$f_{U_{(m+1)}}(x) = \frac{(2m+1)!}{m!m!} x^m (1-x)^m = \frac{1}{B(m+1,m+1)} x^m (1-x)^m$$

that is $U_{(m+1)}$ follows a Beta$(m+1, m+1)$ distribution. Working as in exercise 3 above using MGF's we can see that

$$\begin{aligned}
M_{U_{(m+1)}}(t) &= 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{m+1+r}{2m+2+r} \right) \frac{t^k}{k!} \\
&\overset{m\to\infty}{\longrightarrow} 1 + \sum_{k=1}^{\infty} \left(\prod_{r=0}^{k-1} \frac{1}{2} \right) \frac{t^k}{k!} = 1 + \sum_{k=1}^{\infty} \frac{(t/2)^k}{k!} = e^{t/2}
\end{aligned}$$

that is the MGF of the constant random variable that is $1/2$ with probability one. This result agrees with our intuition that as the sample size keeps increasing, we should expect the median to fall close to $1/2$.

5. Define

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_1^2] \\ \mathbb{E}[X_1^3] \\ \mathbb{E}[X_1^4] \end{pmatrix}$$

and $\boldsymbol{\Omega}$ with $\Omega_{i,j} = \mathrm{cov}(Y_i, Y_j)$. One can show

$$\Omega_{i,j} = \cdots = n\mu_{i+j} - \frac{n(n-1)}{2}\mu_i\mu_j.$$

Then the multivariate CLT states that

$$\sqrt{n}\,(\mathbf{Y_n} - \boldsymbol{\mu}) \xrightarrow{d} \mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{\Omega}).$$

6. We write the expectation

$$\mathbb{E}[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mu$$

and the variance

$$\mathrm{var}[\hat{\mu}] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{var}[X_i] = \frac{\mu}{n},$$

therefore the MSE of $\hat{\mu}$ as an estimator of μ is

$$\mathrm{MSE}(\hat{\mu}, \mu) = \mathrm{var}[\hat{\mu}] + (\mathbb{E}[\hat{\mu}] - \mu)^2 = \frac{\mu}{n}.$$

7. We have that

$$\begin{aligned} f(\mathbf{Y}; p) &= \prod_{i=1}^{n} f(Y_i; p) = \prod_{i=1}^{n} p^{Y_i}(1-p)^{1-Y_i} \\ &= p^{\sum_{i=1}^{n} Y_i}(1-p)^{n-\sum_{i=1}^{n} Y_i}, \end{aligned}$$

therefore

$$\log f(\mathbf{Y}; p) = \left(\sum_{i=1}^{n} Y_i\right)\log p + \left(n - \sum_{i=1}^{n} Y_i\right)\log(1-p)$$

and

$$\frac{\partial^2}{\partial p^2}\log f(\mathbf{Y}; p) = -\frac{\sum_{i=1}^{n} Y_i}{p^2} - \frac{n - \sum_{i=1}^{n} Y_i}{(1-p)^2}$$

so

$$\begin{aligned} \mathcal{I}_n &= -\mathbb{E}\left[\frac{\partial^2}{\partial p^2}\log f(\mathbf{Y}; p)\right] \\ &= \frac{np}{p^2} + \frac{n - np}{(1-p)^2} = \cdots = \frac{n}{p(1-p)} \end{aligned}$$

and we get the Cramer-Rao lower bound $1/\mathcal{I}_n = \frac{p(1-p)}{n}$.