Homework 2

Daniel Hartig

March 1, 2017

Problem 1

For probability function

$$f_{T_n}(t) = (8n - 8n^2 t)\, I_{(1/2n,1/n)}(t)$$

the distribution function over the same domain is obtained by integration

$$F_{T_n}(t) = (8nt - 4n^2t^2 + C)\, I_{(1/2n,1/n)}(t).$$

To be a valid distrubition function, F must equal 0 at the lower bound of the domain, $t = 1/2n$, and equal 1 at the upper bound $t = 1/n$. At theses bounds, F evaluates to

$$\begin{aligned} F_{T_n}(1/2n) &= \frac{8n}{2n} - \frac{4n^2}{(2n)^2} + C \quad &= 3 + C \\ F_{T_n}(1/n) &= \frac{8n}{n} - \frac{4n^2}{n^2} + C \quad &= 4 + C. \end{aligned}$$

Therefore $C = -3$, $F(t) = 0$ for $t < 1/2n$ and $F(t) = 1$ for $t > 1/n$, and

$$F_{T_n}(t) = (8nt - 4n^2t^2 - 3)\, I_{(1/2n,1/n)}(t)$$

meets the requirements of a distrubution function for all n. As n approaches infinity, the range of $I_{(1/2n,1/n)}(t)$ approaches the point 0. We have already defined the $F(t) = 1$ for $t > 1/n$, so we can say that

$$\lim_{n\to\infty} F_{T_n}(t) = 1 \quad \forall t > 0.$$

Thus, the limiting distribution of $F_{T_n}(t)$ approaches $F(t)$, where $F(t)$ is

$$F(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$$

The probability function of the limiting distribution for the sequence $T_1, T_2, ...$ is the derivative of the distribution function, and is

$$f(t) = \begin{cases} 1, & t = 0 \\ 0, & \text{otherwise} \end{cases}$$

Problem 2

For $f_{X_n}(x)$, the moment-generating function is calculated by

$$M_X(t) = \sum_{x\in A} e^{tx} p(x) = e^{t\cdot 0}\left(\frac{n-1}{n}\right) + e^{tn^2}\left(\frac{1}{n}\right) = \frac{1}{n}\left(e^{tn^2} + n - 1\right).$$

To calculate the $E(X_n)$ we need the first t derivative of M_X,

$$M_X'(t) = \frac{d}{dt}\left[\frac{1}{n}\left(e^{tn^2} + n - 1\right)\right] = ne^{tn^2},$$

so that

$$E(X_n) = M'_X(0) = ne^{0\cdot n^2} = n.$$

Thus,

$$\begin{aligned} T_n &= X_n - E(X_n) \\ &= X_n - n \end{aligned}$$

The probility distribution of T_n can then be defined as

$$f_{T_n}(x) = \begin{cases} \frac{1}{n}, & x = n^2 - n \\ \frac{n-1}{n}, & x = -n \\ 0, & \text{otherwise} \end{cases}$$

The cumulative distribution function is

$$F_{T_n}(x) = \begin{cases} 0, & x < -n \\ \frac{n-1}{n}, & -n \le x < n^2 - n \\ 1, & x \ge n^2 - n. \end{cases}$$

As $n \to \infty$, the term $(n-1)/n$ approaches 1, so

$$\lim_{n\to\infty} F_{T_n} = \begin{cases} 0, & x < -n \\ 1, & x \ge -n. \end{cases}$$

However, $-n$ iteself goes to $-\infty$, so F_{T_n} does not converge. Since the distribution function of T_n does not converge, a limiting distrubtion for the sequence $T_1, T_2, ...$ does not exist.

Problem 3

Part a

For the continuious probability function $f_X(x) = 3(1-x)^2\, I_{(0,1)}(x)$, the cumulative distribution function is obtained by integration

$$F_X(x) = \int 3(1-x)^2\, I_{(0,1)}(x)\, dx = (x-1)^3 + C.$$

For the bounds of the probability function 0 and 1, the values of F_X are

$$\begin{aligned} (x-1)^3 + C\Big|_0 &= -1 + C \\ (x-1)^3 + C\Big|_1 &= C. \end{aligned}$$

Set $C = 1$, $F_X = 0$ for $x < 0$ and $F_X = 1$ for $x > 1$, and then

$$F_X(x) = \begin{cases} 0, & x < 0 \\ (x-1)^3 + 1, & 0 < x < 1 \\ 1, & x > 1 \end{cases}$$

is a valid distribution function. Since the maxium possible value of the probability function is $x = 1$, we expect $X_{(n)}$ to converge to 1 as more random variables are added to the sequence. Since we expect the sequence of maxima to converge to 1, we can say that it will converge to a random variable $X = 1$. To apply this to Definition 5.5.1, we solve for

$$\begin{aligned} P\left(\left|X_{(n)} - 1\right| \ge \epsilon\right) &= P\left(X_{(n)} \ge 1 + \epsilon\right) + P\left(X_{(n)} \le 1 - \epsilon\right) \\ &= P\left(X_{(n)} \le 1 - \epsilon\right). \end{aligned}$$

Using the distribution function for X_i, and a change of variables from $0 < x < 1$ to $0 < (1-\epsilon) < 1$, we can say that for all X_i with $1 \le i \le n$,

$$P(X_i \le 1 - \epsilon) = ((1-\epsilon) - 1)^3 + 1 = 1 - \epsilon^3.$$

Since the X_i are independent of each other, the probability that all X_i in the series $X_1, X_2, ..., X_n$ are less than $1 - \epsilon$ is

$$P\left(X_{(n)} < 1 - \epsilon\right) = \left(1 - \epsilon^3\right)^n.$$

This goes to zero for all $\epsilon > 0$, therefore the maximum is proven to converge to 1. We can change variables again to $\epsilon = t/n^{1/3}$ where since $n \geq 1$, $t > 0$ to match ranges. We then get

$$P\left(X_{(n)} \leq 1 - t/n^{1/3}\right) = \left(1 - \left(\frac{t}{n^{1/3}}\right)^3\right)^n = \left(1 - \frac{t^3}{n}\right)^n \to \left(e^{-t^3}\right) I_{(0,\infty)}(t)$$
$$P\left(n^{1/3}\left(1 - X_{(n)}\right) \leq t\right) \to \left(1 - e^{-t^3}\right) I_{(0,\infty)}(t).$$

The left side of this expression is the distribution function for T_n. To find the probability function by Theorem 5.5.12, we must take the t derivative of the distribution function

$$f_{T_n}(t) = \frac{d}{dt}\left[1 - e^{-t^3}\right] = 3\,t^2\,e^{-t^3}\,I_{(0,\infty)}(t).$$

Therefore, as n goes to infinity, the sequence $T_1, T_2, ...$ converges to the above probability function.

Part b

By Theorem 5.5.4, if $X_1, X_2, ...$ converges in probability to X, and h is a continuous function, then $h(X_1), h(X_2), ...$ converges in probability to $h(X)$. Let $h(y) = \sqrt{y}$ and

$$V_n = h(T_n) = \sqrt{T_n}.$$

Since we know that T_n converges in probability to

$$3\,t^2\,e^{-t^3}\,I_{(0,\infty)}(t),$$

then V_n converges to

$$\sqrt{3\,t^2\,e^{-t^3}}\,I_{(0,\infty)}(t) = \sqrt{3}\,t\exp\left(\frac{-t^3}{2}\right)I_{(0,\infty)}(t)$$

Problem 4

The moment generating function of the sum of independent random variables in the sequence $X_1, X_2, ...$ is

$$M_{X_1}(t)M_{X_2}(t)....$$

The moment generating function of a Poisson distribution is $\exp(\lambda(e^t - 1))$. Thus the moment generating function for $T_n = X_1 + ... + X_n$ is

$$\begin{aligned} M_{T_n}(t) &= \prod_{k=1}^{n} \exp\left(k^{-2}\left(e^t - 1\right)\right) \\ &= \exp\left(\left(e^t - 1\right)\sum_{k=1}^{n}\frac{1}{k^2}\right). \end{aligned}$$

This is itself the moment generating function of a Poisson distrubution with

$$\lambda = \sum_{k=1}^{n}\frac{1}{k^2}.$$

As n goes to infinity, the sum of this infinite series is $\pi^2/6$. Therefore, as n goes to infinity, the sequence $T_1, T_2, ...$ converges to a limiting distribution of Poisson$\left(\pi^2/6\right)$, the probability mass function of which is

$$\frac{\left(\frac{\pi^2}{6}\right)^j e^{-\pi^2/6}}{j!} = \frac{\pi^{2j}e^{-\pi^2/6}}{6^j\,j!}$$

Problem 5

For iid randm variables $X_1, X_2, ...,$ the sample mean $\bar{X}$ is $(1/n)\sum X_i$. In otherwords,

$$\sum_{i=1}^{n} X_i = n\bar{X}.$$

Since X_i is a Bernoulli variable, the only possible outcomes are 0 and 1. $0^2 = 0$ and $1^2 = 1$, $X_i^2 = X_i$, and

$$\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i = n\bar{X}.$$

We can re-write the formula for T_n as

$$T_n = \sqrt{n}\left(\frac{4n\bar{X} - 2n}{n\bar{X}}\right) = \sqrt{n}\left(\frac{\bar{X} - \frac{1}{2}}{\frac{1}{4}\bar{X}}\right).$$

By the Central Limit Theorem, this function has a limiting standard normal distribution with $\mu = 1/2$ and $\sigma = (1/4)\bar{X}$. We can find $\bar{X}$ by the Law of Large numbers

$$\lim_{n\to\infty} \bar{X}_n = \mu = p = \frac{1}{2}$$

as given in the problem statement. Therefore, as $n \to \infty$, the sequence $T_1, T_2, ...$ converges to a limiting distribution of Normal$(1/2, 1/8)$.

Problem 6

Part a

X is a random variable with $E(X) = \theta$. Let $g(X) = 1/\sqrt{X}$ be an estimator for $1/\sqrt{\theta}$; $g'(X) = -1/2\, x^{-3/2}$. The first order Taylor approximation to the mean of $1/\sqrt{X_n}$ is

$$\begin{aligned} g(X) &= g(\theta) + g'(\theta)(X - \theta) \\ \frac{1}{\sqrt{X_n}} &= \frac{1}{\sqrt{\theta}} - \frac{1}{2}\theta^{-\frac{3}{2}}(X_n - \theta) \\ E\left[\frac{1}{\sqrt{X_n}}\right] &= \frac{1}{\sqrt{\theta}} + -\frac{1}{2}\theta^{-\frac{3}{2}}\, E\left[X_n - \theta\right] \end{aligned}$$

Since θ is the mean of X, $E[X - \theta] = 0$, and so

$$E\left[\frac{1}{\sqrt{X}}\right] = \frac{1}{\sqrt{\theta}}$$

Part b

Since $g''(x) = 3/4x^{-5/2}$; the second order approximation to the mean of $1/\sqrt{X_n}$ is

$$\begin{aligned} g(X) &= g(\theta) + g'(\theta)(X - \theta) + \frac{1}{2}g''(\theta)(X_n - \theta)^2 \\ E\left[\frac{1}{\sqrt{X}}\right] &= \frac{1}{\sqrt{\theta}} + \frac{3}{4}\theta^{-\frac{5}{2}} E\left[(X_n - \theta)^2\right] \end{aligned}$$

Now θ is the mean of X, $E\left[(X - \theta)^2\right]$ is the definition of the variance of X, which is θ^3. Therefore

$$E\left[\frac{1}{\sqrt{X_n}}\right] = \frac{1}{\sqrt{\theta}} + \frac{3}{4}\sqrt{\theta} = \frac{1 + \frac{3}{4}\theta}{\sqrt{\theta}}$$

Part c

The first order approximation to the variance of $1/\sqrt{X_n}$ is

$$\begin{aligned} \text{Var}\left[\frac{1}{\sqrt{X_n}}\right] &= \left(-\frac{1}{2}\theta^{-\frac{3}{2}}\right)^2 \text{Var}\left[X_n\right] \\ &= \frac{1}{4}\frac{1}{\theta^3}\theta^3 = \frac{1}{4} \end{aligned}$$

Problem 7

To generate a sufficient statistic, we must factorize the joint probability density function $f_X(\mathbf{x}|\theta)$ into two terms $g(T(\mathbf{x})|\theta)$ and $h(\mathbf{x})$ as

$$\begin{aligned} f_X(x) &= \prod_{k=1}^{n} (\theta+1) x_k^{\theta} I_{(0,1)}(x) \\ &= (\theta+1)^n \exp\left(\theta \sum_{k=1}^{n} \log x_k\right) I_{(0,1)}(x). \end{aligned}$$

Therefore

$$h(\mathbf{x}) = I_{(0,1)}(x)$$

and

$$g(T(\mathbf{x})|\theta) = (\theta+1)^n \exp\left(\theta \sum_{k=1}^{n} \log x_k\right),$$

so

$$T(\mathbf{x}) = \sum_{k=1}^{n} \log x_k$$

is a sufficient statistic for θ for every x in the sample space.