2011 Theory I Problem #1

1. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. real random variables with $\mathbb{E}X_1 = 0$. Let N be a Poisson random variable with parameter $\lambda \geq 0$ and is independent of $X_1, X_2, \ldots$ For each integer $m \geq 0$, let $\overline{X}_m = m^{-1}\sum_{i=1}^m X_i$, where we define $\overline{X}_0 = 0$.

   (a) Assume $\sigma^2 = \mathbb{E}X_1^2 < \infty$ and do the following:

      (i) Show that $\text{Var}[\overline{X}_N] \leq \sigma^2\left[\mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \geq \lambda^{1/3})}{\lambda}\right]$

      It can be shown that the claim does not hold for all λ. We will instead prove a weaker claim which is sufficiently sharp to obtain the desired results in the following sections. We wish to show that

$$\text{Var}[\overline{X}_N] \leq \sigma^2\left[\mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \geq \lambda^{1/3})}{\lambda^{1/3}}\right]$$

      First we observe that $\mathbb{E}[X_m] = 0,\ m = 0, 1, \ldots$ and that

$$\text{Var}[\overline{X}_m] = \begin{cases} 0, & m = 0 \\ \sigma^2/m, & m = 1, 2, \ldots \end{cases}$$

      It follows that

$$\begin{aligned}
\text{Var}[\overline{X}_N] &= \mathbb{E}\Big(\text{Var}[\overline{X}_N|N]\Big) + \text{Var}\Big(\mathbb{E}[\overline{X}_N|N]\Big) \\
&= \mathbb{E}\Big(\text{Var}[\overline{X}_N|N]\Big) + 0 \\
&= \sum_{m=0}^{\infty} \text{Var}[\overline{X}_m]\,\mathbb{P}(N = m) \\
&= \sum_{m=1}^{\infty} \text{Var}[\overline{X}_m]\,\mathbb{P}(N = m) \\
&= \sum_{m=1}^{\infty} \frac{\sigma^2}{m}\mathbb{P}(N = m) \\
&= \sigma^2 \sum_{m=1}^{\infty} \frac{1}{m}\mathbb{P}(N = m)
\end{aligned}$$

      Thus it suffices to show that

$$\sum_{m=1}^{\infty} \frac{1}{m}\mathbb{P}(N = m) \leq \mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \geq \lambda^{1/3})}{\lambda^{1/3}}$$

We consider two cases. First suppose that $\lambda \le 1$. Then

$$\begin{aligned}
\sum_{m=1}^{\infty} \frac{1}{m} \mathbb{P}(N=m) &\le \sum_{m=1}^{\infty} \frac{1}{\lambda^{1/3}} \mathbb{P}(N=m) \\
&= \frac{\mathbb{P}(N \ge 1)}{\lambda^{1/3}} \\
&\le P(N=0) + \frac{\mathbb{P}(N \ge 1)}{\lambda^{1/3}} \\
&= \mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \ge \lambda^{1/3})}{\lambda^{1/3}}
\end{aligned}$$

Now suppose that $\lambda > 1$ and let M_λ be the largest integer less than $\lambda^{1/3}$. Then

$$\begin{aligned}
\sum_{m=1}^{\infty} \frac{1}{m} \mathbb{P}(N=m) &= \sum_{m=1}^{M_\lambda} \frac{1}{m} \mathbb{P}(N=m) + \sum_{M_\lambda+1}^{\infty} \frac{1}{m} \mathbb{P}(N=m) \\
&\le \sum_{m=1}^{M_\lambda} \frac{1}{1} \mathbb{P}(N=m) + \sum_{M_\lambda+1}^{\infty} \frac{1}{\lambda^{1/3}} \mathbb{P}(N=m) \\
&= \mathbb{P}(1 \le N \le M_\lambda) + \frac{1}{\lambda^{1/3}} \mathbb{P}(N \ge M_\lambda + 1) \\
&\le \mathbb{P}(0 \le N \le M_\lambda) + \frac{1}{\lambda^{1/3}} \mathbb{P}(N \ge M_\lambda + 1) \\
&= \mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \ge \lambda^{1/3})}{\lambda^{1/3}}
\end{aligned}$$

(ii) Show that $\mathbb{P}(N < \lambda^{1/3}) \to 0$ as $\lambda \to \infty$. Hint: use Chebychev's inequality.

By Chebychev's inequality we have

$$\begin{aligned}
\mathbb{P}(N < \lambda^{1/3}) &< \mathbb{P}\left(N < \lambda^{1/3} \ \text{ or } \ N > \lambda^{5/3}\right) \\
&= \mathbb{P}\left(|N-\lambda| > \lambda^{2/3}\right) \\
&\le \frac{\mathrm{Var}[N]}{(\lambda^{2/3})^2} \\
&= \frac{\lambda}{\lambda^{4/3}} \to 0 \quad \text{as} \quad \lambda \to \infty
\end{aligned}$$

(iii) Show that $\lim_{\lambda\to\infty} \mathbb{P}\left(|\overline{X}_N| \ge \epsilon\right) = 0$ for every $\epsilon > 0$.

By Chebychev's inequality and our results in (1a.i) and (1a.ii) we have

$$\begin{aligned}
\mathbb{P}\left(|\overline{X}_N| \ge \epsilon\right) &\le \frac{\mathrm{Var}[\overline{X}_N]}{\epsilon^2} \\
&\le \frac{\sigma^2}{\epsilon^2}\left[\mathbb{P}(N < \lambda^{1/3}) + \frac{\mathbb{P}(N \ge \lambda^{1/3})}{\lambda^{1/3}}\right] \\
&\le \frac{\sigma^2}{\epsilon^2}\left[\mathbb{P}(N < \lambda^{1/3}) + \frac{1}{\lambda^{1/3}}\right] \to 0 \quad \text{as} \quad \lambda \to \infty
\end{aligned}$$

(b) Let $\psi(t)$ be the characteristic function of a standard normal random variable, and define $Z_m = m^{1/2}\overline{X}_m/\sigma$. Continue to assume $\sigma^2 < \infty$. Do the following:

(i) Show that for any real t,

$$\left| \mathbb{E}[e^{itZ_N}] - \psi(t) \right| \le 2\,\mathbb{P}(N < \lambda^{1/3}) + \max_{m \ge \lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right|$$

Recall that by Euler's formula $|e^{it}| = 1$. Further recall that for arbitrary integrable random variable W it holds that $\left|\mathbb{E}[W]\right| \le \mathbb{E}|W|$. Use of these facts and multiple applications of the triangle inequality yields

$$\begin{aligned}
\left| \mathbb{E}[e^{itZ_N}] - \psi(t) \right| &= \left| \mathbb{E}[e^{itZ_N} I_{N<\lambda^{1/3}}] + \mathbb{E}[e^{itZ_N} I_{N\ge\lambda^{1/3}}] - \psi(t) \right| \\
&= \left| \mathbb{E}[e^{itZ_N} I_{N<\lambda^{1/3}}] + \mathbb{E}_N\Big(\mathbb{E}_{Z_N|N}[e^{itZ_N}\,|N] I_{N\ge\lambda^{1/3}}\Big) - \psi(t) \right| \\
&\le \left| \mathbb{E}[e^{itZ_N} I_{N<\lambda^{1/3}}] \right| + \left| \mathbb{E}_N\Big(\mathbb{E}_{Z_N|N}[e^{itZ_N}\,|N] I_{N\ge\lambda^{1/3}}\Big) - \psi(t) \right| \\
&\le \mathbb{E}\Big[|e^{itZ_N}| I_{N<\lambda^{1/3}}\Big] + \max_{m\ge\lambda^{1/3}} \left| \mathbb{E}_N\Big(\mathbb{E}_{Z_m}[e^{itZ_m}] I_{N\ge\lambda^{1/3}}\Big) - \psi(t) \right| \\
&= \mathbb{E}\Big[1 \cdot I_{N<\lambda^{1/3}}\Big] + \max_{m\ge\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}]\mathbb{P}(N \ge \lambda^{1/3}) - \psi(t) \right| \\
&= \mathbb{P}(N<\lambda^{1/3}) + \max_{m\ge\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}]\Big(1 - \mathbb{P}(N<\lambda^{1/3})\Big) - \psi(t) \right| \\
&\le \mathbb{P}(N<\lambda^{1/3}) + \max_{m\ge\lambda^{1/3}} \left\{ \left| -\mathbb{E}[e^{itZ_m}]\mathbb{P}(N<\lambda^{1/3}) \right| + \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| \right\} \\
&\le \mathbb{P}(N<\lambda^{1/3}) + \max_{m\ge\lambda^{1/3}} \left\{ \mathbb{E}|e^{itZ_m}|\mathbb{P}(N<\lambda^{1/3}) + \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| \right\} \\
&= \mathbb{P}(N<\lambda^{1/3}) + \max_{m\ge\lambda^{1/3}} \left\{ 1 \cdot \mathbb{P}(N<\lambda^{1/3}) + \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| \right\} \\
&= 2\,\mathbb{P}(N<\lambda^{1/3}) + \max_{m\ge\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right|
\end{aligned}$$

(ii) Show that for any real t, $\left|\mathbb{E}[e^{itZ_N}] - \psi(t)\right| \to 0$ as $\lambda \to \infty$.

From our results in (1a.ii) and (1b.i) it suffices to show that

$$\max_{m\ge\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| \to 0 \quad \text{as} \quad \lambda \to \infty$$

From the central limit theorem $Z_m \xrightarrow{L} N(0,1)$, so that by the continuity theorem $\mathbb{E}[e^{itZ_m}] \to \psi(t)$. Then for every $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that for all $m \ge N_\epsilon$,

$$\left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| < \epsilon$$

Thus for the above ϵ, choosing $\lambda > N_\epsilon^3$ implies that

$$\max_{m\ge\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| < \epsilon$$

and since ϵ is arbitrary we obtain that

$$\lim_{\lambda\to\infty} \max_{m\geq\lambda^{1/3}} \left| \mathbb{E}[e^{itZ_m}] - \psi(t) \right| = 0$$

(c) Now do not assume $\sigma^2 < \infty$. Do the following:

(i) Show that for each $\epsilon > 0$

$$\mathbb{P}\Big(\big|\overline{X}_N\big| \geq \epsilon\Big) \leq \mathbb{P}\big(N < \lambda^{1/3}\big) + \mathbb{P}\left(\max_{m\geq\lambda^{1/3}} \big|\overline{X}_m\big| \geq \epsilon\right)$$

$$\begin{aligned}
\mathbb{P}\left(|\overline{X}_N| \geq \epsilon\right) &= \mathbb{E}\left[I\left(|\overline{X}_N| \geq \epsilon\right)\right] \\
&= \mathbb{E}\left\{I\left(|\overline{X}_N| \geq \epsilon,\, N < \lambda^{1/3}\right) + I\left(|\overline{X}_N| \geq \epsilon,\, N \geq \lambda^{1/3}\right)\right\} \\
&= \mathbb{P}\left(|\overline{X}_N| \geq \epsilon,\, N < \lambda^{1/3}\right) + \mathbb{P}\left(|\overline{X}_N| \geq \epsilon,\, N \geq \lambda^{1/3}\right) \\
&\leq \mathbb{P}\left(|\overline{X}_N| \geq \epsilon,\, N < \lambda^{1/3}\right) + \mathbb{P}\left(\max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon,\, N \geq \lambda^{1/3}\right) \\
&\leq \mathbb{P}\left(N < \lambda^{1/3}\right) + \mathbb{P}\left(\max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right)
\end{aligned}$$

(ii) Show that $\overline{X}_N \xrightarrow{p} 0$ as $\lambda \to \infty$. Hint: use the strong law of large numbers.

From (1a.ii) and (1c.i) it suffices to show that

$$\mathbb{P}\left(\max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right) \to 0 \quad \text{as} \quad \lambda \to \infty$$

From the SLLN we obtain that $\lim_{m\to\infty} \overline{X}_m = 0$. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be the probability space upon which $\overline{X}_m$ resides, and let $\mathscr{N} \subset \Omega$ denote the set upon which $\overline{X}_m \to 0$. Consider $\epsilon > 0$. Then for $\omega \in \mathscr{N}$, there exists a $N_{\epsilon,\omega} \in \mathbb{N}$ such that for all $m \geq N_{\epsilon,\omega}$ it holds that $|\overline{X}_m(\omega)| < \epsilon$. It follows that for such an ω we find that

$$\lim_{\lambda\to\infty} \max_{m\geq\lambda^{1/3}} |\overline{X}_m(\omega)| = 0$$

and consequently

$$\mathbb{P}\left(\lim_{\lambda\to\infty} \max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right) = 0$$

Now let g be the function which identically maps each element of Ω to the value 1. Then

$$I\left(\max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right) \leq g \quad \text{and} \quad \int g\, dP = 1 < \infty$$

so that we may invoke the dominated convergence theorem yielding

$$0 = \mathbb{P}\left(\lim_{\lambda\to\infty} \max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right) = \lim_{\lambda\to\infty} \mathbb{P}\left(\max_{m\geq\lambda^{1/3}} |\overline{X}_m| \geq \epsilon\right)$$

2011 Theory I Problem #2

2. (a) Let X be a random variable and let ν be a parameter of interest in the distribution of X. Suppose that $T(X)$ is an unbiased estimator of ν. Show that any unbiased estimator of ν is of the form $T(X) - U(X)$, where $\mathbb{E}[U(X)] = 0$.

Consider some other unbiased estimator of ν, say $W(X)$. Define $U(X) = T(X) - W(X)$. Then

$$\mathbb{E}[U(X)] = \mathbb{E}[T(X) - W(X)] = \mathbb{E}[T(X)] - \mathbb{E}[W(X)] = \nu - \nu = 0$$

Furthermore,

$$W(X) = T(X) - U(X)$$

In the sequel, let X be a discrete random variable with $\mathbb{P}(X = -1) = p$ and $\mathbb{P}(X = k) = (1-p)^2 p^k$, $k = 0, 1, 2, \ldots$, where $p \in (0, 1)$ is unknown.

(b) Show that $\mathbb{E}[U(X)] = 0$ if and only if $U(k) = ak$ for all $k = -1, 0, 1, \ldots$ and some a.

We wish to find a U such that $\mathbb{E}[U(X)] = 0$. To begin with, we express $\mathbb{E}[U(X)]$ as a polynomial in p where each coefficent is a function of U. Then in order for $\mathbb{E}[U(X)]$ to be uniformly 0 for each $p \in (0, 1)$ we will need to find a choice of U such that each coefficent is 0. We have

$$\begin{aligned}
0 &\overset{set}{=} \mathbb{E}[U(X)] \\
&= pU(-1) + \sum_{x=0}^{\infty} U(x)(1-p)^2 p^x \\
&= pU(-1) + \sum_{x=0}^{\infty} U(x)[1 - 2p + p^2] p^x \\
&= pU(-1) + \sum_{x=0}^{\infty} U(x) p^x - 2\sum_{x=0}^{\infty} U(x) p^{x+1} + \sum_{x=0}^{\infty} U(x) p^{x+2} \\
&= pU(-1) + \sum_{x=-2}^{\infty} U(x+2) p^{x+2} - 2\sum_{x=-1}^{\infty} U(x+1) p^{x+2} + \sum_{x=0}^{\infty} U(x) p^{x+2} \\
&= pU(-1) + \left[U(0) + \sum_{x=-1}^{\infty} U(x+2) p^{x+2}\right] - 2\sum_{x=-1}^{\infty} U(x+1) p^{x+2} + \left[\sum_{x=0}^{\infty} U(x) p^{x+2} - pU(-1)\right] \\
&= U(0) + \sum_{x=-1}^{\infty} \Big[U(x) - 2U(x+1) + U(x+2)\Big] p^{x+2}
\end{aligned}$$

Now we wish to argue that $U(0)$ must be 0 for $\mathbb{E}[U(X)]$ to uniformly equal 0. Suppose for a moment that $U(0) \neq 0$. Then

$$-U(0) = \sum_{x=-1}^{\infty} \Big[U(x) - 2U(x+1) + U(x+2)\Big] p^{x+2}$$

for all $p \in (0,1)$. However, this implies that $U(X)$ must be a function of p, which it may not be as it an estimator; thus we conclude that $U(0) = 0$.

Thus we may reduce the problem to finding $U(\cdot)$ that satisfies

$$U(x) - 2U(x+1) + U(x+2) = 2, \; x = -1, 0, 1, \ldots$$

Incorporating the fact that $U(0)$ must equal 0, we wish to solve the following system of equations:

$$\begin{aligned}
0 &= U(-1) + 0 + \; U(1) \\
0 &= \quad\quad\quad 0 - 2U(1) + \; U(2) \\
0 &= \quad\quad\quad\quad\quad U(1) - 2U(2) + \; U(3) \\
0 &= \quad\quad\quad\quad\quad\quad\quad U(2) - 2U(3) + U(4) \\
&\vdots \quad\quad\quad\quad\quad\quad \ddots \quad \ddots \quad \ddots
\end{aligned}$$

By summing equalities {1,2}, {1,2,3}, {1,2,3,4},... we find that

$$\begin{aligned}
0 &= U(-1) + U(1) \\
0 &= U(-1) - U(1) + U(2) \\
0 &= U(-1) \quad\quad - U(2) + U(3) \\
&\vdots \quad\quad\quad\quad \ddots \quad \ddots
\end{aligned}$$

or equivalently (after some rearrangement and substitution)

$$\begin{cases} U(0) = 0 \\ -U(-1) = U(1) \\ U(1) = U(x) - U(x-1), \quad x = 2, 3, \ldots \end{cases}$$

We conclude that $U(k) = ak$ solves the system of equations for any choice of a (where a is our $U(1)$).

(c) Using the results in (a) and (b), show that $I(X = 0)$ is the unique admissible estimator under squared error loss amongst all unbiased estimators of $(1-p)^2$, where $I(\cdot)$ is the indicator function.

Let $T(X) = I(X = 0)$. We observe that

$$\mathbb{E}[T(X)] = \sum_{x=-1}^{\infty} I(x=0)(1-p)^2 p^x = (1-p)^2$$

$$\mathrm{Var}[T(X)] = \mathbb{E}\left([T(X)]^2\right) - \left(\mathbb{E}[T(X)]\right)^2 = \mathbb{E}[T(X)] - \left(\mathbb{E}[T(X)]\right)^2 = (1-p)^2 - (1-p)^4$$

and note that T is unbiased for $(1-p)^2$.

Now consider some other unbiased estimator or $(1-p)^2$, say $W(X)$. Since T is unbiased we have from

(2a) and (2b) that

$$\begin{aligned}
&\mathbb{E}\left\{\left[W(X)-(1-p)^2\right]^2\right\}\\
&\quad=\mathbb{E}\left\{\left[T(X)-U(X)-(1-p)^2\right]^2\right\}\\
&\quad=\mathbb{E}\left\{\left[T(X)-U(X)\right]^2\right\}-2(1-p)^2\,\mathbb{E}\left[T(X)-U(X)\right]+(1-p)^4\\
&\quad=\left\{\left[T(-1)-U(-1)\right]^2p+\sum_{x=0}^{\infty}\left[T(x)-U(x)\right]^2(1-p)^2p^x\right\}\\
&\qquad\qquad\qquad\qquad-2(1-p)^2\left(\mathbb{E}\left[T(X)\right]-\mathbb{E}\left[U(X)\right]\right)+(1-p)^4\\
&\quad=\left\{\left[I(-1=0)-(-a)\right]^2p+\sum_{x=0}^{\infty}\left[T(x)-U(x)\right]^2(1-p)^2p^x\right\}\\
&\qquad\qquad\qquad\qquad-2(1-p)^2\left((1-p)^2-0\right)+(1-p)^4\\
&\quad=a^2p+\sum_{x=0}^{\infty}\left[T(x)-U(x)\right]^2(1-p)^2p^x-(1-p)^4\\
&\quad\geq a^2p-(1-p)^4=\mathrm{Var}\left[T(X)\right]=\mathbb{E}\left\{\left[T(X)-(1-p)^2\right]^2\right\}
\end{aligned}$$

(d) Show that no unique admissible estimator exists for p under squared error loss amongst unbiased estimators for p.

Let $T(X)=I(=-1)$. Then $\mathbb{E}[T(X)]=P(X=-1)=p$ so that T is an unbiased estimator of p. Consider another unbiased estimator of p, say $W(X)$. Then from (2a) and (2b) we have

$$\begin{aligned}
&\mathbb{E}\left\{\left[W(X)-p\right]^2\right\}\\
&\quad=\mathbb{E}\left\{\left[W(X)-U(X)-p\right]^2\right\}\\
&\quad=\mathbb{E}\left\{\left[W(X)-U(X)\right]^2\right\}-2p\mathbb{E}\left[W(X)-U(X)\right]+p^2\\
&\quad=\left\{\left[I(-1=-1)-(-a)\right]^2p+\sum_{x=0}^{\infty}\left[I(x=-1)-ax\right]^2(1-p)^2p^x\right\}\\
&\qquad\qquad\qquad\qquad\quad-2p\Big(\mathbb{E}\left[W(X)\right]-\mathbb{E}\left[U(X)\right]\Big)+p^2\\
&\quad=(1-a)^2p+a^2\sum_{x=0}^{\infty}(1-p)^2p^{x+2}-2p(p-0)+p^2\\
&\quad=(1-a)^2p+a^2\sum_{x=0}^{\infty}(1-p)^2-p^2
\end{aligned}$$

To find an admissible estimator we need to minimize the above expression with respect to a. We have

$$
\begin{aligned}
0 &\overset{set}{=} \frac{d}{da}\left\{(1-a)^2p + a^2\sum_{x=0}^{\infty}(1-p)^2 - p^2\right\} \\
&= -2(1-a)p + 2a\sum_{x=0}^{\infty}x^2(1-p)^2p^x \\
&= -2p + 2a\left[p + \sum_{x=0}^{\infty}x^2(1-p)^2p^x\right] \\
\implies \quad a &= \left[1 + \sum_{x=0}^{\infty}x^2(1-p)^2p^{x-1}\right]^{-1}
\end{aligned}
$$

Since $(1-a)^2p + a^2\sum_{x=0}^{\infty}(1-p)^2 - p^2$ is quadratic in a, the choice of a above is the unique minimizer for a fixed p; note that a changes with p. Now an admissible estimator d is unique if and only if for any other admissible estimator d^* it holds that $d = d^*$. Suppose now there exists a unique admissible estimator d_{opt}. It is clear that we can find a p such that $d^* \neq d_{\text{opt}}$ where

$$
d^* = \arg\min_{d} \mathbb{E}_p\left\{\left[d(X) - p\right]^2\right\}
$$

Since d^* is the unique minimizer for this choice of p we see that d^* is not inadmissible; hence it is admissible. But then d_{opt} is not unique and a contradiction is reached.

(e) Prove whether there exists unbiased estimators of p^{-1}. If so, then determine whether a unique admissible estimator exists under squared error loss amongst unbiased estimators for p^{-1}.

Let $g(X)$ be an unbiased estimator of p^{-1}. Then

$$
p^{-1} = \mathbb{E}[g(X)]
$$

so that

$$
\begin{aligned}
1 &= p\,\mathbb{E}[g(X)] \\
&= p\left[pg(-1) + \sum_{x=0}^{\infty}g(x)(1-p)^2p^x\right] \\
&= p^2g(-1) + \sum_{x=0}^{\infty}g(x)[1-2p+p^2]p^{x+1} \\
&= p^2g(-1) + \sum_{x=0}^{\infty}g(x)p^{x+1} - 2\sum_{x=0}^{\infty}g(x)p^{x+2} + \sum_{x=0}^{\infty}g(x)p^{x+3} \\
&= p^2g(-1) + \left[g(0)p + g(1)p^2 + \sum_{x=2}^{\infty}g(x)p^{x+1}\right] \\
&\qquad - \left[2g(0)p^2 + 2\sum_{x=1}^{\infty}g(x)p^{x+2}\right] + \sum_{x=0}^{\infty}g(x)p^{x+3} \\
&= p^2g(-1) + \left[g(0)p + g(1)p^2 + \sum_{x=0}^{\infty}g(x+2)p^{x+3}\right]
\end{aligned}
$$

$$-\left[2g(0)p^2 + 2\sum_{x=0}^{\infty} g(x+1)p^{x+3}\right] + \sum_{x=0}^{\infty} g(x)p^{x+3}$$

$$= g(0)p + \left[g(-1) - 2g(0) + g(1)\right]p^2 + \sum_{x=0}^{\infty}\left[g(x) - 2g(x+1) + g(x+2)\right]p^{x+3}$$

$$= g(0)p + \sum_{x=-1}^{\infty}\left[g(x) - 2g(x+1) + g(x+2)\right]p^{x+3}$$

From this form we can see that for g to satisfy the above equation for all values of p it must be a function of p. Since an estimator may only be a function of the data we conclude that no unbiased estimator of p^{-1} exists.

2011 Theory I Problem #3

3. Consider a sequence of numbers $x_1, x_2, \ldots$ and place vertical lines before x_1 and between x_j and x_{j+1} whenever $x_j > x_{j+1}$. We say that the runs are the segments between pairs of lines. Thus each run is an incerasing segment of the sequence $x_1, x_2, \ldots$

   Suppose that $X_1, X_2, \ldots$ are independent and identically distributed uniform(0,1) random variables and that we are interested in the lengths of the successive runs. Let L_j denote the length of the j^{th} run.

   (a) Compute $\mathbb{P}(L_1 \geq m),\ m = 1, 2, \ldots$

   We claim that

$$\mathbb{P}\left(L_1 \geq m\right) = \mathbb{P}\left(X_1 \leq \cdots \leq X_m\right) = \frac{1}{m!}$$

   The first equality is immediately apparent and we proceed with proof-by-induction to show the second. Consider first

$$\mathbb{P}(X_1 < X_2) = \mathbb{E}\left[\mathbb{P}\left(X_1 \leq X_2 \,\middle|\, X_2\right)\right] = \mathbb{E}\left[F(X_2)\right] = \mathbb{E}X_2 = \frac{1}{2}$$

   Next, by the induction hypothesis

$$\begin{aligned}
\mathbb{P}\left(X_1 \leq \cdots \leq X_m\right) &= \mathbb{E}\left[\mathbb{P}\left(X_1 \leq \cdots \leq X_m \,\middle|\, X_m\right)\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(X_1 \leq \cdots \leq X_{m-1},\ \max(X_1, \ldots, X_{m-1}) \leq X_m \,\middle|\, X_m\right)\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(X_1 \leq \cdots \leq X_{m-1} \,\middle|\, X_m\right)\mathbb{P}\left(\max(X_1, \ldots, X_{m-1}) \leq X_m \,\middle|\, X_m\right)\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(X_1 \leq \cdots \leq X_{m-1} \,\middle|\, X_m\right)\right]\mathbb{E}\left[\mathbb{P}\left(\max(X_1, \ldots, X_{m-1}) \leq X_m \,\middle|\, X_m\right)\right] \\
&= \mathbb{P}\left(X_1 \leq \cdots \leq X_{m-1}\right)\mathbb{E}\left[\mathbb{P}\left(X_1 \leq X_m, \ldots, X_{m-1} \leq X_m \,\middle|\, X_m\right)\right] \\
&= \frac{1}{(m-1)!}\,\mathbb{E}\left[\prod_{i=1}^{m-1}\mathbb{P}\left(X_i \leq X_m \,\middle|\, X_m\right)\right] \\
&= \frac{1}{(m-1)!}\,\mathbb{E}\left[\prod_{i=1}^{m-1}F(X_m)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(m-1)!}\ \mathbb{E}\left[\prod_{i=1}^{m-1} X_m\right] \\
&= \frac{1}{(m-1)!}\ \mathbb{E}\left[X_m^{m-1}\right] \\
&= \frac{1}{(m-1)!}\ \int_0^1 x^{m-1}\,dx \\
&= \frac{1}{(m-1)!}\ \left.\frac{x^m}{m}\right|_0^1 \\
&= \frac{1}{m!}
\end{aligned}
$$

(b) Suppose that we know that the j^{th} run starts with the value x. Compute $\mathbb{P}(L_j \geq m \,|\, x)$.

$$
\begin{aligned}
\mathbb{P}(L_j \geq m \,|\, x) &= \mathbb{P}\Big(x \leq X_1 \leq \ldots \leq X_{m-1}\Big) \\
&= \mathbb{E}\Big\{\ \mathbb{P}\Big(X_1 \leq \ldots \leq X_{m-1},\ \min(X_1,\ldots,X_{m-1}) > x \ \Big|\ \min(X_1,\ldots,X_{m-1}) > x\Big)\ \Big\} \\
&= \mathbb{E}\Big\{\ \mathbb{P}\Big(X_1 \leq \ldots \leq X_{m-1}\Big)\, I\Big(\min(X_1,\ldots,X_{m-1}) > x\Big)\Big\} \\
&= \mathbb{E}\left[\frac{1}{(m-1)!}\, I\Big(X_1 > x,\ldots,X_{m-1} > x\Big)\right] \\
&= \frac{1}{(m-1)!}\mathbb{E}\left[\prod_{i=1}^{m-1} I\Big(X_i > x\Big)\right] \\
&= \frac{1}{(m-1)!}\prod_{i=1}^{m-1}\mathbb{E}\left[I\Big(X_i > x\Big)\right] \\
&= \frac{1}{(m-1)!}\prod_{i=1}^{m-1}\mathbb{P}\Big(X_i > x\Big) \\
&= \frac{1}{(m-1)!}\Big[\mathbb{P}(X_1 > x)\Big]^{m-1} \\
&= \frac{(1-x)^{m-1}}{(m-1)!}
\end{aligned}
$$

(c) Let I_j denote the initial value of the j^{th} run. Show that $p_n(y|x)$ the probability density that the $(n+1)^{\text{st}}$ run has $I_{n+1} = y$ given that the n^{th} run has just begun with $I_n = x$, equals e^{1-x} if $y < x$ and $e^{1-x} - e^{y-x}$ if $y > x$.

The $(n+1)^{\text{th}}$ run will start with value y conditional on the n^{th} run starting with value x if for some positive integer m

(i) the next $(m-1)$ values in the sequence are in increasing order and are all greater than x
(ii) the m^{th} value must equal y
(iii) the $(m-1)^{\text{th}}$ value is greater than y

Notice that we may replace condition (iii) with $\max(X_k,\ldots,X_{k+m-1}) > y$ for a run starting at X_k without changing the overall statement, which will prove helpful in what follows. We start with a few results and then piece everything together at the end.

$$
\begin{aligned}
&\mathbb{P}\Big(I_{n+1} \in (y,\, y+\Delta),\; L_n = m \;\Big|\; I_n = x\Big) \\
&\quad = \mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1},\, X_m \in (y,\, y+\Delta),\; \max(X_1,\ldots,X_{m-1}) > y\Big) \\
&\quad = \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big)\, \mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1},\; \max(X_1,\ldots,X_{m-1}) > y\Big) \\
&\quad = \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big)\, \mathbb{E}\Big\{\mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1}, \\
&\qquad\qquad\qquad \max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big)\Big\} \\
&\quad = \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big)\, \mathbb{E}\Big\{\mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1} \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big)\Big\} \\
&\qquad\qquad\qquad \times \mathbb{E}\Big\{\mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big)\Big\} \\
&\quad = \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big)\, \mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1}\Big) \\
&\qquad\qquad\qquad \times \mathbb{E}\Big\{\mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big)\Big\}
\end{aligned}
$$

Now when $y < x$,

$$\mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) = 1$$

and when $y > x$,

$$
\begin{aligned}
&\mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&\quad = 1 - \mathbb{P}\Big(X_1 \le y,\ldots,X_{m-1} \le y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&\quad = 1 - \Big[\mathbb{P}\Big(X_1 \le y \;\Big|\; X_1 > x\Big)\Big]^{m-1} \\
&\quad = 1 - \left[\frac{\mathbb{P}(x < X_1 \le y)}{\mathbb{P}(X_1 > x)}\right]^{m-1} \\
&\quad = 1 - \left(\frac{y-x}{1-x}\right)^{m-1}
\end{aligned}
$$

Also,

$$f_{X_m}(y) = \lim_{\Delta \searrow 0} \frac{1}{\Delta}\, \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big) = \lim_{\Delta \searrow 0} \frac{1}{\Delta}\, \Delta = 1, \qquad \text{for all } m$$

Then

$$
\begin{aligned}
p(y|x) &= \lim_{\Delta\searrow 0} \frac{1}{\Delta} \mathbb{P}\Big(I_{n+1} \in (y,\, y+\Delta) \;\Big|\; I_n = x\Big) \\
&= \lim_{\Delta\searrow 0} \frac{1}{\Delta} \sum_{m=1}^{\infty} \mathbb{P}\Big(I_{n+1} \in (y,\, y+\Delta),\; L_n = m \;\Big|\; I_n = x\Big) \\
&= \lim_{\Delta\searrow 0} \frac{1}{\Delta} \sum_{m=1}^{\infty} \mathbb{P}\Big(X_m \in (y,\, y+\Delta)\Big)\, \mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1}\Big) \\
&\qquad\qquad\qquad \times \mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&= \lim_{\Delta\searrow 0} \frac{1}{\Delta} \mathbb{P}\Big(X_1 \in (y,\, y+\Delta)\Big) \sum_{m=1}^{\infty} \mathbb{P}\Big(x \le X_1 \le \cdots \le X_{m-1}\Big) \\
&\qquad\qquad\qquad \times \mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&= \sum_{m=1}^{\infty} \frac{(1-x)^{m-1}}{(m-1)!}\, \mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big)
\end{aligned}
$$

Thus by the Taylor series expansion for e^a, when $y < x$,

$$
\begin{aligned}
p(y|x) &= \sum_{m=1}^{\infty} \frac{(1-x)^{m-1}}{(m-1)!}\, \mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&= \sum_{m=1}^{\infty} \frac{(1-x)^{m-1}}{(m-1)!} \\
&= e^{1-x}
\end{aligned}
$$

and for $y > x$,

$$
\begin{aligned}
p(y|x) &= \sum_{m=1}^{\infty} \frac{(1-x)^{m-1}}{(m-1)!}\, \mathbb{P}\Big(\max(X_1,\ldots,X_{m-1}) > y \;\Big|\; X_i > x,\, i = 1,\ldots,m-1\Big) \\
&= \sum_{m=1}^{\infty} \frac{(1-x)^{m-1}}{(m-1)!} \left[1 - \left(\frac{y-x}{1-x}\right)^{m-1}\right] \\
&= \sum_{m=1}^{\infty} \left[\frac{(1-x)^{m-1}}{(m-1)!} - \frac{(y-x)^{m-1}}{(m-1)!}\right] \\
&= e^{1-x} - e^{y-x}
\end{aligned}
$$

(d) Demonstrate that $\pi(y)$, the probability density function for I_n as $n \to \infty$, satisfies $\pi(y) = 2(1-y)$, $0 < y < 1$. You may do this by verifying the continuous state equilibrium equations for discrete time Markov chains: $\pi(y) = \int_0^1 \pi(x)\, p(y|x)\, dx$.

Wish to verify that

$$
2(1-y) = \pi(y) = \int_0^1 \pi(x)\, p(y|x)\, dx = \int_0^1 2(1-x)\, p(y|x)\, dx
$$

$$
\iff \quad 1 - y = \int_0^1 (1-x)\, p(y|x)\, dx
$$

Now,

$$\int_0^1 (1-x)\,p(y|x)\,dx = \int_0^1 (1-x)\Big[(e^{1-x}-e^{y-x})\,I(y>x)+e^{1-x}I(y<x)\Big]dx$$

$$= \int_0^y (1-x)(e^{1-x}-e^{y-x})\,dx + \int_y^1 (1-x)\,e^{1-x}dx$$

$$= \int_0^1 (1-x)\,e^{1-x}dx - \int_y^1 (1-x)\,e^{y-x}dx$$

$$= \int_0^1 (1-x)\,e^{1-x}dx - e^{-(1-y)}\int_0^y (1-x)\,e^{1-x}dx$$

$$= \int_0^1 ze^z dz - e^{-(1-y)}\int_{1-y}^1 ze^z dz$$

$$= (ze^z-e^z)\Big|_0^1 - e^{-(1-y)}(ze^z-e^z)\Big|_{1-y}^1$$

$$= \Big[(e-e)-(0-1)\Big]-e^{-(1-y)}\Big[(1-1)-\Big((1-y)\,e^{1-y}-e^{1-y}\Big)\Big]$$

$$= 1+e^{-(1-y)}\Big[(1-y)\,e^{1-y}-e^{1-y}\Big]$$

$$= 1+(1-y)-1$$

$$= 1-y$$

(e) Find $\lim_{n\to\infty}\mathbb{P}(L_n \geq m)$.

Let π_n denote the density of the initial value of the n^{th} run. Now

$$\mathbb{P}(L_n \geq m) = \mathbb{E}\Big[\mathbb{P}\Big(L_n \geq m \,\Big|\, I_n\Big)\Big]$$

$$= \int_0^1 \mathbb{P}\Big(L_n \geq m \,\Big|\, I_n\Big)\,\pi_n(x)\,dx$$

$$= \int_0^1 \frac{(1-x)^{m-1}}{(m-1)!}\,\pi_n(x)\,dx$$

Next, we observe that

$$\frac{(1-x)^{m-1}}{(m-1)!} \leq 1 \qquad \text{for all } x \in (0,1)$$

so that for any density $p(x)$,

$$\int_0^1 \frac{(1-x)^{m-1}}{(m-1)!}\,p(x)\,dx \leq \int_0^1 p(x)\,dx = 1$$

Then the limiting distribution of the length of the n^{th} run is given by

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(L_n \geq m) &= \lim_{n\to\infty} \int_0^1 \frac{(1-x)^{m-1}}{(m-1)!}\, \pi_n(x)\, dx \\
&= \int_0^1 \lim_{n\to\infty} \frac{(1-x)^{m-1}}{(m-1)!}\, \pi_n(x)\, dx \\
&= \int_0^1 \frac{(1-x)^{m-1}}{(m-1)!}\, \pi(x)\, dx \\
&= \int_0^1 \frac{(1-x)^{m-1}}{(m-1)!}\, 2(1-x)\, dx \\
&= \frac{2}{(m-1)!} \int_0^1 (1-x)^m dx \\
&= \frac{2}{(m-1)!} \int_0^1 x^m dx \\
&= \frac{2}{(m-1)!} \left. \frac{x^{m+1}}{m+1} \right|_{x=0}^{x=1} \\
&= \frac{2}{(m+1)(m-1)!}
\end{aligned}$$

where the interchange of the limit and integral is justified by the dominated convergence theorem.

(f) What is the average length of a run as $n \to \infty$, that is, what is $\lim_{n\to\infty} \mathbb{E}[L_n]$?

First we observe that

$$\begin{aligned}
\sum_{m=1}^{\infty} \mathbb{P}(L_n \geq m) &= \sum_{m=1}^{\infty} \sum_{k=m}^{\infty} \mathbb{P}(L_n = k) \\
&= \sum_{k=1}^{\infty} \sum_{m=1}^{k} \mathbb{P}(L_n = k) \\
&= \sum_{k=1}^{\infty} k\, \mathbb{P}(L_n = k) \\
&= \mathbb{E}[L_n]
\end{aligned}$$

where interchange of the integrals is justified by Fubini's theorem for nonnegative functions (sometimes called Tonelli's theorem). Then

$$\begin{aligned}
\mathbb{E}[L_n] &= \mathbb{E}\Big(\mathbb{E}\big[L_n \,|\, I_n\big]\Big) \\
&= \mathbb{E}\left(\sum_{m=1}^{\infty} \mathbb{P}\big[L_n \geq m \,|\, I_n\big]\right) \\
&= \mathbb{E}\left(\sum_{m=1}^{\infty} \frac{(1-I_n)^{m-1}}{(m-1)!}\right)
\end{aligned}$$

$$= \mathbb{E}\left[e^{1-I_n}\right]$$

so that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[L_n] &= \lim_{n\to\infty} \mathbb{E}\left[e^{1-I_n}\right] \\
&= \lim_{n\to\infty} \int_0^1 e^{1-x} \pi_n(x)\, dx \\
&= \int_0^1 \lim_{n\to\infty} e^{1-x} \pi_n(x)\, dx \\
&= \int_0^1 e^{1-x} \pi(x)\, dx \\
&= \int_0^1 e^{1-x} 2(1-x)\, dx \\
&= 2\int_0^1 z e^z dz \\
&= 2(ze^z - e^z)\big|_0^1 \\
&= 2
\end{aligned}$$

where interchange of the limit and the integral is justified via the dominated convergence theorem in a similar manner as was done in (e).

An alternative solution can be obtained as follows. We have

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[L_n] &= \lim_{n\to\infty} \sum_{m=1}^{\infty} \mathbb{P}(L_n \geq m) \\
&= \sum_{m=1}^{\infty} \lim_{n\to\infty} \mathbb{P}(L_n \geq m) \\
&= \sum_{m=1}^{\infty} \frac{2}{(m+1)(m-1)!} \\
&= 2\sum_{m=1}^{\infty} \left(\frac{1}{m!} - \frac{1}{(m+1)!}\right) \\
&= 2\Big[(e-1)-(e-2)\Big] \\
&= 2
\end{aligned}$$