Homework for 1/8

1. [§4-6] Let $X$ be a continuous random variable with probability density function $f(x) = 2x, 0 \leq x \leq 1$.

   (a) Find $E[X]$.
   (b) Find $E[X^2]$.
   (c) Find $\text{Var}(X)$.

(a) We have
\[
E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_0^1 x \cdot 2x \, dx = \left. \left( \frac{2x^3}{3} \right) \right|_0^1 = \frac{2}{3}.
\]

(b) We have
\[
E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) \, dx = \int_0^1 x^2 \cdot 2x \, dx = \left. \left( \frac{x^4}{2} \right) \right|_0^1 = \frac{1}{2}.
\]

(c) We have
\[
\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \left( \frac{2}{3} \right)^2 = \frac{1}{18}.
\]

2. [§4-31] Let $X$ be uniformly distributed on the interval $[1, 2]$. Find $E[1/X]$.

Is $E[1/X] = 1/E[X]$?

Since $X$ is uniformly distributed on $[1, 2]$, the pdf of $X$ is
\[
f(x) = \begin{cases} 
1, & 1 \leq x \leq 2 \\
0, & \text{otherwise}.
\end{cases}
\]

On one hand, we have
\[
E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_1^2 x \cdot 1 \, dx = \left. \left( \frac{x^2}{2} \right) \right|_1^2 = \frac{3}{2}.
\]

On the other hand, we have
\[
E \left[ \frac{1}{X} \right] = \int_{-\infty}^{\infty} \frac{1}{x} f(x) \, dx = \int_1^2 \frac{1}{x} \cdot 1 \, dx = (\ln x) \bigg|_1^2 = \ln 2 - \ln 1.
\]

It follows immediately that $E[1/X] \neq 1/E[X]$. 
3. §4-42 Let \( X \) be an exponential random variable with standard deviation \( \sigma \). Find \( P(|X - E[X]| > k\sigma) \) for \( k = 2, 3, 4 \), and compare the results to the bounds from Chebyshev’s inequality.

Since \( X \) is exponentially distributed and has variance \( \sigma^2 \), the pdf of \( X \) is

\[
f(x) = \begin{cases} \frac{1}{\sigma} e^{-x/\sigma}, & x > 0 \\ 0, & \text{otherwise} \end{cases}
\]

It follows immediately that \( E[X] = \sigma \). Thus we have

\[
P(|X - E[X]| > k\sigma) = P(X > (k+1)\sigma \text{ or } X < (1-k)\sigma)
\]

\[
= \int_{(k+1)\sigma}^\infty \frac{1}{\sigma} e^{-x/\sigma} \, dx \quad \text{we only consider } k = 2, 3, 4
\]

\[
= \int_{k+1}^\infty e^{-y} \, dy \quad \text{(change of variable } y = x/\sigma) 
\]

\[
= e^{-(k+1)}. 
\]

Therefore, we have

\[
P(|X - E[X]| > 2\sigma) = e^{-3} = 0.0498, \quad P(|X - E[X]| > 3\sigma) = e^{-4} = 0.0183, 
\]

\[
P(|X - E[X]| > 4\sigma) = e^{-5} = 0.0067. 
\]

On the other hand, the Chebyshev’s inequality gives us

\[
P(|X - E[X]| > k\sigma) \leq \frac{\text{Var}[X]}{k^2\sigma^2} = \frac{1}{k^2}. 
\]

In particular, we have

\[
P(|X - E[X]| > 2\sigma) \leq \frac{1}{2^2} = 0.25, \quad P(|X - E[X]| > 3\sigma) \leq \frac{1}{3^2} = 0.1111, 
\]

\[
P(|X - E[X]| > 4\sigma) \leq \frac{1}{4^2} = 0.0625. 
\]

4. §4-80 Let \( X \) be a continuous random variable with density function \( f(x) = 2x, 0 \leq x \leq 1 \). Find the moment-generating function of \( X \), \( M(t) \), and verify that \( E[X] = M'(0) \) and that \( E[X^2] = M''(0) \).
The mgf of $X$ is

$$M(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx = \int_{0}^{1} e^{tx} \cdot 2x \, dx$$

$\quad = 2 \left[ \frac{e^{tx}}{t} \left|_{0}^{1} \right. - \frac{1}{t} \int_{0}^{1} e^{tx} \, dx \right]$ 

$\quad = 2 \left[ \frac{e^t}{t} - \frac{1}{t^2} (e^t - 1) \right]$ (integration by parts) 

$\quad = 2 \left( \frac{e^t}{t} - \frac{e^t}{t^2} + \frac{1}{t^2} \right)$.

Note that by L'Hospital rule, we have

$$\lim_{t \to 0} M(t) = 2 \lim_{t \to 0} \frac{te^t - e^t + 1}{t^2} = 2 \lim_{t \to 0} \frac{te^t + e^t - e^t}{2t} = \lim_{t \to 0} e^t = 1.$$ 

Furthermore, we have

$$M'(t) = 2 \left( \frac{e^t}{t} - \frac{2e^t}{t^2} + \frac{2e^t - 2}{t^3} \right), \quad \text{and}$$

$$\lim_{t \to 0} M'(t) = 2 \lim_{t \to 0} \frac{t^2e^t - 2te^t + 2e^t - 2}{t^3}$$

$\quad = 2 \lim_{t \to 0} \frac{t^2e^t + te^t - 2e^t + 2e^t}{3t^2}$

$\quad = 2 \lim_{t \to 0} e^t = \frac{2}{3},$ 

and

$$M''(t) = 2 \left( \frac{e^t}{t} - \frac{3e^t}{t^2} + \frac{6e^t}{t^3} - \frac{6e^t}{t^4} + \frac{6}{t^4} \right), \quad \text{and}$$

$$\lim_{t \to 0} M''(t) = 2 \lim_{t \to 0} \frac{t^3e^t - 3t^2e^t + 6te^t - 6e^t + 6}{t^4}$$

$\quad = 2 \lim_{t \to 0} \frac{t^3e^t + 3t^2e^t - 3t^2e^t - 6te^t + 6te^t + 6e^t - 6e^t}{4t^3}$

$\quad = \frac{1}{2} \lim_{t \to 0} e^t = \frac{1}{2}.$

Combining with the first problem (§4-6), we see that $\mathbb{E}[X] = M'(0)$ and $\mathbb{E}[X^2] = M''(0).$
5. [§4-91] Use the mgf to show that if $X$ follows an exponential distribution, $cX(c > 0)$ does also.

Let the pdf of $X$ be

$$f(X) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}.$$  

Then the mgf of $X$ is

$$M_X(t) = \int_0^\infty e^{tx} \cdot \lambda e^{-\lambda x} \, dx = \frac{\lambda}{\lambda - t} \int_0^\infty (\lambda - t)e^{-(\lambda - t)} \, dx = \frac{\lambda}{\lambda - t},$$

for all $t < \lambda$. (The trick here is to identify a pdf for $\text{Exp}(\lambda - t)$). Let $Y = cX$. Then the mgf of $Y$ is

$$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(cX)}] = \mathbb{E}[e^{(ct)X}] = M_X(ct) = \frac{\lambda}{\lambda - ct} = \frac{\lambda/c}{\lambda/c - t},$$

for all $t$ with $ct < \lambda$, that is, $t < \lambda/c$. Compare $M_Y(t)$ to $M_X(t)$, we see that $Y$ is also exponentially distributed and the parameter is $\lambda/c$.

Note that the result

$$M_{cX}(t) = M_X(ct)$$

is generally true as long as the mgf are defined (for example, we need to make sure $ct < \lambda$). More generally, we have

$$M_{aX+b}(t) = e^{bt}M_X(t).$$
Homework for 1/10

1. [§5-16] Suppose that $X_1, \ldots, X_{20}$ are independent random variables with density functions
\[ f(x) = 2x, \quad 0 \leq x \leq 1 \]
Let $S = X_1 + \cdots + X_{20}$. Use the central limit theorem to approximate $\mathbb{P}(S \leq 10)$.

From Problem 1 of Homeowrk 1/8, we have
\[ \mu = \mathbb{E}[X] = \frac{2}{3}, \quad \text{and} \quad \sigma^2 = \text{Var}[X] = \frac{1}{18} \]
Therefore, according the central limit theorem, we have
\[
\mathbb{P}(S \leq 10) = \mathbb{P}\left( \frac{S - n\mu}{\sigma/\sqrt{n}} \leq \frac{10 - n\mu}{\sigma/\sqrt{n}} \right) = \mathbb{P}\left( \frac{S - 20 \cdot \frac{2}{3}}{\sqrt{\frac{1}{18} \cdot \sqrt{20}}} \leq \frac{10 - 20 \cdot \frac{2}{3}}{\sqrt{\frac{1}{18} \cdot \sqrt{20}}} \right)
\]
\[
\approx \mathbb{P}(Z \leq -3.16) = 0.0008,
\]
where $Z \sim N(0, 1)$.

2. [§5-17] Suppose that a measurement has mean $\mu$ and variance $\sigma^2 = 25$. Let $X$ be the average of $n$ such independent measurements. How large should $n$ be so that $\mathbb{P}(|X - \mu| < 1) = .95$?

By applying central limit theorem, we have
\[
0.95 = \mathbb{P}(|X - \mu| < 1) = \mathbb{P}\left( \frac{|X - \mu|}{\sigma/\sqrt{n}} < \frac{1}{\sigma/\sqrt{n}} \right)
\]
\[
\approx \mathbb{P}\left( |Z| < \frac{1}{\sqrt{n}} \right) = \mathbb{P}\left( -\frac{1}{\sqrt{n}} < Z < \frac{1}{\sqrt{n}} \right),
\]
where $Z \sim N(0, 1)$. Since $0.95 = \mathbb{P}(-1.96 < Z < 1.96)$, we have
\[
\frac{1}{\sqrt{n}} = 1.96,
\]
which implies that
\[
n = (5 \cdot 1.96)^2 = 96.04 \approx 96.
\]
3. Suppose that a company ships packages that are variable in weight, with an average weight of 15 lb and a standard deviation of 10. Assuming that the packages come from a large number of different customers so that it is reasonable to model their weights as independent random variables, find the probability that 100 packages will have a total weight exceeding 1700 lb.

Let $X_1, X_2, \ldots, X_{100}$ denote the weight of these 100 packages, and $S_{100} = X_1 + \cdots + X_{100}$ the total weight. Since $\mu = \mathbb{E}[X_i] = 15$ and $\sigma^2 = \text{Var}[X_i] = 10^2$ for $1 \leq i \leq 100$, the central limit theorem gives

$$
\mathbb{P}(S_{100} > 1700) = \mathbb{P}
\left(
\frac{S_{100} - 100\mu}{\sigma \sqrt{100}} > \frac{1700 - 100\mu}{\sigma \sqrt{100}}
\right)
= \mathbb{P}
\left(
\frac{S_{100} - 100 \cdot 15}{10 \cdot \sqrt{100}} > \frac{1700 - 100 \cdot 15}{10 \cdot \sqrt{100}}
\right)
\approx \mathbb{P}(Z > 2) = 0.0228.
$$