MATH2033 Mathematical Statistics

Suggested Solution to Assignment 1

1. (a) Given: U is uniform on $[0,1]$

$$a = 0$$
$$b = 1$$

The probability density function of a uniform distribution is the reciprocal of the difference of the boundaries, on the interval between the boundaries (0 elsewhere):

$$f_U(x) = \frac{1}{b-a} = \frac{1}{1-0} = 1$$

Obviously, $\sqrt{U} \in [0,1]$. Consider $u \in [0,1]$. Then cumulative distribution function of $\sqrt{U}$ at u is

$$F_{\sqrt{U}}(u) = P(\sqrt{U} \leq u)$$

Square each side of the inequality and use that u is nonnegative (because U is only defined for nonnegative values):

$$F_{\sqrt{U}}(u) = P\left(U \leq u^2\right)$$

Use the definition of the cumulative distribution function of U :

$$F_{\sqrt{U}}(u) = F_U\left(u^2\right)$$

The density is the derivative of the cumulative distribution function:

$$f_{\sqrt{U}}(u) = 2u f_U\left(u^2\right)$$

Use the probability density function of U to obtain

$$f_{\sqrt{U}}(u) = 2u f_U\left(u^2\right) = 2u(1) = 2u, \quad u \in [0,1],$$

and for $u \notin [0,1]$, $f_{\sqrt{U}}(u) = 0$.

(b) Given: U is uniform on $[-1,1]$

$$a = -1$$
$$b = 1$$

The probability density function of a uniform distribution is the reciprocal of the difference of the boundaries, on the interval between the boundaries (0 elsewhere):

$$f_U(x) = \frac{1}{b-a} = \frac{1}{1-(-1)} = \frac{1}{2}$$

Obviously, $U^2 \in [0,1]$. Consider $u \in [0,1]$. The cumulative distribution function of U^2 at u is

$$F_{U^2}(u) = P\left(U^2 \leq u\right)$$

Take the square root of each side of the inequality (Note: u has to be nonnegative):

$$F_{U^2}(u) = P(-\sqrt{u} \leq U \leq \sqrt{u})$$

Use the definition of the cumulative distribution function of U :

$$F_{U^2}(u) = F_U(\sqrt{u}) - F_U(-\sqrt{u})$$

The density is the derivative of the cumulative distribution function:

$$f_{U^2}(u) = \frac{1}{2\sqrt{u}} f_U(\sqrt{u}) - \frac{1}{2\sqrt{u}} f_U(-\sqrt{u})$$

Use the probability density function of U: for $u \in [0,1]$,

$$\begin{aligned}
f_{U^2}(u) &= \frac{1}{2\sqrt{u}} f_U(\sqrt{u}) - \frac{1}{2\sqrt{u}} f_U(-\sqrt{u}) \\
&= \frac{1}{2\sqrt{u}} \frac{1}{2} - \frac{1}{2(-\sqrt{u})} \frac{1}{2} \\
&= \frac{1}{4\sqrt{u}} + \frac{1}{4\sqrt{u}} \\
&= \frac{2}{4\sqrt{u}} \\
&= \frac{1}{2\sqrt{u}}
\end{aligned}$$

and for $u \notin [0,1]$, $f_{U^2}(u) = 0$.

2. (a) For each fixed $x_2 \in (0,1)$, $f_{1|2}\left(x_1 \mid x_2\right)$ should be a pdf itself, thus

$$c_1 \int_0^{x_2} x_1 / x_2^2 dx_1 = \frac{c_1}{2} = 1 \Rightarrow c_1 = 2.$$

   Similarly, $\int_0^1 f_2\left(x_2\right) dx_2 = 1$ and thus $c_2 = 5$.

   (b) $10 x_1 x_2^2, 0 < x_1 < x_2 < 1$; zero elsewhere

   (c) $\int_{1/4}^{1/2} 2x_1/(5/8)^2 dx = \frac{64}{25}\left(\frac{1}{4} - \frac{1}{16}\right) = \frac{12}{25}$.

   (d) $\int_{1/4}^{1/2} \int_{x_1}^{1} 10 x_1 x_2^2 dx_2 dx_1 = \int_{1/4}^{1/2} \frac{10}{3} x_1 \left(1 - x_1^3\right) dx_1 = \frac{449}{1536}$.

3. (a) The marginal pdf of X_2 is

$$f_2\left(x_2\right) = \int_0^{x_2} 21 x_1^2 x_2^3 dx_1 = 7 x_2^6, \quad 0 < x_2 < 1.$$

   Therefore,

$$f_{1|2}\left(x_1 \mid x_2\right) = 21 x_1^2 x_2^3 / 7 x_2^6 = 3 x_1^2 / x_2^3, \quad 0 < x_1 < x_2.$$

$$E\left(X_1 \mid X_2 = x_2\right) = \int_0^{x_2} x_1 \left(3 x_1^2 / x_2^3\right) dx_1 = \frac{3}{4} x_2.$$

Moreover,

$$
\begin{aligned}
E\left[X_1^2 \mid X_2=x_2\right] &= \int_0^{x_2} x_1^2 \cdot \left(3x_1^2/x_2^3\right) dx_1 \\
&= \frac{3}{x_2^3} \cdot \frac{1}{5} \cdot x_2^5 = \frac{3}{5} x_2^2. \\
\Rightarrow Var\left[X_1 \mid X_2=x_2\right] &= E\left[X_1^2 \mid X_2=x_2\right] - \left(E\left[X_1 \mid X_2=x_2\right]\right)^2 \\
&= \frac{3}{5}x_2^2 - \frac{9}{16}x_2^2 = \frac{48-45}{80}x_2^2 = \frac{3}{80}x_2^2.
\end{aligned}
$$

(b) According to (a), we know that

$$
Y = E\left[X_1 \mid X_2\right] = \frac{3}{4}X_2.
$$

So

$$
\begin{aligned}
F_Y(y) &= P\left(\frac{3}{4}X_2 \le y\right) = \int_0^{4y/3} 7x_2^6 dx_2 = \left(\frac{4y}{3}\right)^7, \quad 0 < y < \frac{3}{4} \\
f_Y(y) &= \begin{cases} 7\left(\frac{4}{3}\right)^7 y^6 & 0 < y < \frac{3}{4} \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}
$$

(c) We have:

$$
\begin{aligned}
E[Y] &= \int_0^{\frac{3}{4}} y \cdot 7 \cdot \left(\frac{4}{3}\right)^7 y^6 dy \\
&= \frac{1}{8} \cdot 7 \cdot \left(\frac{4}{3}\right)^7 \cdot \left(\frac{3}{4}\right)^8 \\
&= \frac{7}{8} \cdot \frac{3}{4} = \frac{21}{32}. \\
\mathrm{Var}(Y) &= E\left[Y^2\right] - (E[Y])^2. \\
&= \int_0^{\frac{3}{4}} y^2 \cdot 7 \cdot \left(\frac{4}{3}\right)^7 y^6 dy - \left(\frac{21}{32}\right)^2 \\
&= \frac{1}{9} \cdot 7 \cdot \left(\frac{4}{3}\right)^7 \cdot \left(\frac{3}{4}\right)^9 - \left(\frac{21}{32}\right)^2 \\
&= \frac{7}{16} \cdot \frac{9}{16} - \frac{441}{1024} = \frac{7}{1024}.
\end{aligned}
$$

$$
\begin{aligned}
E\left(X_1\right) &= \frac{21}{32} = E[Y]. \\
\mathrm{Var}\left(X_1\right) &= \frac{553}{15360} > \frac{7}{1024} = \mathrm{Var}(Y).
\end{aligned}
$$

4. (a) $P\left(X_1 < 1/2\right) = \int_0^{1/2} 4x^3 dx = x^4\Big|_0^{1/2} = 1/16$

(b) This asks for the probability of exactly one success in a binomial experiment with $n=4$ and $p=1/16$, so the probability is $\binom{4}{1}(1/16)^1(15/16)^3 = 0.206.$

(c) $f_{X_1,X_2,X_3,X_4}(x_1,x_2,x_3,x_4) = \prod_{j=1}^{4} 4x_j^3 = 256\,(x_1x_2x_3x_4)^3\,, 0 \le x_1,x_2,x_3,x_4 \le 1$

(d) $f_{X_2,X_3}(x_2,x_3) = 16x_2^3x_3^3,\ 0 \le x_2,x_3 \le 1.$
When $x_2 < 0$ or $x_3 < 0$, $F_{X_2,X_3}(x_2,x_3) = 0.$
When $0 \le x_2,x_3 < 1$, $F_{X_2,X_3}(x_2,x_3) = \int_0^{x_3}\int_0^{x_2}(4s^3)(4t^3)\,dsdt = \int_0^{x_2} 4s^2ds\int_0^{x_3} 4t^3dt = x_2^4x_3^4$
When $x_2 \ge 1,\ 0 \le x_3 < 1$, $F_{X_2,X_3}(x_2,x_3) = \int_0^{x_3}\int_0^{1}(4s^3)(4t^3)\,dsdt = x_3^4$
When $0 \le x_2 < 1,\ x_3 \ge 1$, $F_{X_2,X_3}(x_2,x_3) = \int_0^{1}\int_0^{x_2}(4s^3)(4t^3)\,dsdt = x_2^4$
When $x_2 \ge 1,\ x_3 \ge 1$, $F_{X_2,X_3}(x_2,x_3) = \int_0^{1}\int_0^{1}(4s^3)(4t^3)\,dsdt = 1$

Thus, $F_{X_2,X_3}(x_2,x_3) = \begin{cases} 0,\ x_2 < 0 \text{ or } x_3 < 0; \\ x_2^4x_3^4,\ 0 \le x_2,x_3 < 1; \\ x_3^4,\ x_2 \ge 1,\ 0 \le x_3 < 1; \\ x_2^4,\ 0 \le x_2 < 1,\ x_3 \ge 1; \\ 1,\ x_2 \ge 1,\ x_3 \ge 1. \end{cases}$

5. We are given that $X_1, X_2, \ldots, X_{20}$ are independent identically distributed random variables with the density function

$$f(x) = 2x, \quad x \in [0,1].$$

A random variable S is defined as

$$S = \sum_{i=1}^{20} X_i,$$

and we need to find the approximate probability $P(S \le 10)$. Usually, $n = 20$ is not large enough to use the Central Limit Theorem (i.e. the approximation will not be satisfying), but we can still use it to obtain some sort of approximation. Remember that the Central Limit Theorem states that

$$\frac{S - 20 \cdot E(X_1)}{\sqrt{20 \cdot \text{Var}(X_1)}} \overset{D}{\approx} N(0,1),$$

where D means the convergence in distribution. The expected value of X_1 (and all other X_i' s) is

$$E(X_1) = \int_0^1 x \cdot f(x)dx = \int_0^1 x \cdot 2xdx = 2 \cdot \int_0^1 x^2dx = 2 \cdot \left(\frac{x^3}{3}\right)\Bigg|_0^1 = 2 \cdot \frac{1}{3} = \frac{2}{3},$$

its second moment is

$$E\left(X_1^2\right) = \int_0^1 x^2 \cdot f(x)dx = \int_0^1 x^2 \cdot 2xdx = 2 \cdot \int_0^1 x^3dx = 2 \cdot \left(\frac{x^4}{4}\right)\Bigg|_0^1 = 2 \cdot \frac{1}{4} = \frac{1}{2},$$

which means that the variance of X_1 is

$$\text{Var}(X_1) = E\left(X_1^2\right) - [E(X_1)]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

Therefore,

$$\frac{S - 20 \cdot \frac{2}{3}}{\sqrt{20 \cdot \frac{1}{18}}} = \frac{S - \frac{40}{3}}{\sqrt{\frac{10}{9}}} \overset{D}{\approx} N(0,1),$$

Using the above approximation we derived using the Central Limit Theorem, we have that the desired probability is

$$\begin{aligned} P(S \leq 10) &= P\left(\frac{S - \frac{40}{3}}{\sqrt{\frac{10}{9}}} \leq \frac{10 - \frac{40}{3}}{\sqrt{\frac{10}{9}}}\right) = P\left(\frac{S - \frac{40}{3}}{\sqrt{\frac{10}{9}}} \leq -3.16\right) \\ &\approx \Phi(-3.16) = 1 - \Phi(3.16) = 1 - 0.9992 \\ &= 0.0008. \end{aligned}$$