1. Let $X$ be a continuous random variable with a density function which is symmetric about 0. Show that $E(X) = 0.$

Solution: Since $f(x)$ is symmetric about 0, we have $f(-x) = f(x).$ This means $xf(x)$ is an odd function. It immediately follows that

$$E(X) = \int_{-\infty}^{\infty} xf(x) \, dx = 0.$$ 

2. Let $X_1, X_2, \ldots, X_{100}$ be independent random variables with the common density $f(x) = 2 - 2x, 0 \leq x \leq 1.$ Let $S = X_1 + X_2 \ldots X_{100}.$ Use the Central Limit Theorem to estimate $P(S \leq 35).$

Solution: We have for each $i$:

$$E(X_i) = \int_0^1 x(2 - 2x) \, dx = \frac{1}{3},$$
$$E(X_i^2) = \int_0^1 x^2(2 - 2x) \, dx = \frac{1}{6},$$
and

$$Var(X_i) = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$ 

Thus $E(S) = \frac{100}{3}$ and the standard deviation of $S$ is $\frac{10}{3\sqrt{2}}.$ By the Central Limit Theorem, $S$ is approximately normal, and we get

$$P(S \leq 35) \approx P\left(Z \leq \frac{35 - \frac{100}{3}}{\frac{10}{3\sqrt{2}}} \right) = P(Z \leq .7071) = .7602.$$ 

3. Compute the moment generating function of a geometric random variable, and use it to compute the mean and the variance.

Solution: Let $X$ be geometric with parameter $p$:

$$M(t) = \sum_{x=1}^{\infty} e^{tx} p(1 - p)^{x-1} = \frac{p e^t}{1 - (1 - p)e^t}.$$ 

Differentiating and evaluating at 0 gives $E(X) = 1/p$ and $Var(X) = \frac{1 - p}{p^2}.$
4. In a certain community, 4 percent of all adults over the age of 50 have tuberculosis (T.B.). A health service in this community correctly diagnoses 98 percent of all persons with T.B. as having the disease, and incorrectly diagnoses 3 percent of all persons without T.B. as having the disease. Find the probabilities that,

(a) the community health service will diagnose an adult over 50 as having T.B.,

Solution: Let PD = positive diagnosis, ND = negative diagnosis, TB = has TB, NTB = doesn’t have TB. Then we have \( P(PD|TB) = .98 \), \( P(PD|NTB) = .03 \), \( P(TB) = .04 \), and \( P(NTB) = .96 \). and

\[
\]

(b) a person over 50 diagnosed by the health service as having T.B. actually has the disease.

\[
P(TB|PD) = \frac{P(TB \cap PD)}{P(PD)} = \frac{P(PD|TB)P(TB)}{P(PD)} = \frac{(.98)(.04)}{.068} = .5765.
\]

5. Suppose the joint probability density of \( X \) and \( Y \) is given by

\[
f(x, y) = \begin{cases} 
24y(1 - x - y) & \text{for } x > 0, y > 0, x + y < 1, \\
0 & \text{elsewhere.}
\end{cases}
\]

(a) find the marginal density of \( X \).

Solution: the joint density is zero except on the triangle bounded by the \( x \)-axis, the \( y \)-axis, and the line \( x + y = 1 \).

\[
f_X(x) = \int_{0}^{1-x} 24y(1 - x - y) \, dy = 24\left[\frac{y^2}{2}(1 - x) - \frac{y^3}{3}\right]_{0}^{1-x} = 4(1 - x)^3, \quad 0 \leq x \leq 1.
\]

(b) find the marginal density of \( Y \).

\[
f_Y(y) = \int_{0}^{1-y} 24y(1 - x - y) \, dx = 24y\left[x - \frac{x^2}{2} - yx\right]_{0}^{1-y} = 12y^3 - 24y^2 + 12y, \quad 0 \leq y \leq 1.
\]

(c) determine if the two variables are independent.

From parts a) and b) above we see that \( f_{X,Y}(x, y) \neq f_X(x)f_Y(y) \) so they are not independent.

6. Let \( X \) have variance \( \sigma^2 \) and let \( m_i = E(X^i) \) denote the \( i \)th moment. The skewness of the random variable \( X \) is defined to be

\[
\text{skw}(X) = E((X - m_1)^3)/\sigma^3.
\]
(a) Show that

\[ \text{skw}(X) = \frac{m_3 - 3m_1m_2 + 2m_1^3}{\sigma^3}, \]

Solution: \((X - m_1)^3 = X^3 - 3X^2m_1 + 3Xm_1^2 - m_1^3\) so

\[
E((X - m_1)^3) = E(X^3) - 3E(X^2m_1) + 3E(Xm_1^2) - E(m_1^3) = m_3 - 3m_1m_2 + 3m_1m_1^2 - m_1^3 = m_3 - 3m_1m_2 + 2m_1^3
\]

and the result follows.

(b) Compute the skewness of an exponential variable with parameter \(\lambda\) and show that it doesn’t depend on \(\lambda\). (The best way to compute the moments is from the moment generating function.)

If \(X\) is exponential with parameter \(\lambda\), then \(M_X(t) = \frac{\lambda}{\lambda - t}\). From this we get the moments \(m_1 = \frac{1}{\lambda}\), \(m_2 = \frac{2}{\lambda^2}\), and \(m_3 = \frac{6}{\lambda^3}\). Substituting into the above formula we get \(\text{skw}(X) = 2\).

7. Suppose you play a series of 100 independent games. If you win a game, you win 4 dollars. If you lose a game, you lose 4 dollars. The chances of winning each game is \(1/2\). Use the central limit theorem to estimate the chances that you will win more than 50 dollars.

Solution. Let \(X_i\) be a random variable which is 4 with probability \(1/2\) and \(-4\) with probability \(1/2\). Then for each \(i\):

\[
E(X_i) = 0,
\]

\[
E(X_i^2) = \text{Var}(X_i) = 16.
\]

Letting \(S = X_1 + \ldots X_{100}\) Thus \(E(S) = 0\), the variance of \(S\) is 1600, and the standard deviation of \(S\) is 40. By the Central Limit Theorem, \(S\) is approximately normal, and we get

\[
P(S > 50) \approx P(Z > \frac{50 - 0}{40}) = P(Z > 1.25) = .1057.
\]

8. Let \(X\) and \(Y\) be continuous random variables, having joint probability density function

\[
f(x, y) = \begin{cases} 
24xy & \text{for } 0 < x < 1, 0 < y < 1, 0 < x + y < 1, \\
0 & \text{elsewhere}.
\end{cases}
\]

Find the probability density of \(Z = X + Y\).
Solution: There are various ways to do this. One way is to first compute the cdf of $Z$, and then differentiate to get the pdf:

$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = \int \int_R f(x, y) \, dx \, dy$$

where $R$ is the triangular region bounded by the $x$-axis, the $y$-axis, and the line $x + y = z$. We get

$$\int \int_R f(x, y) \, dx \, dy = \int_0^z \int_0^{z-y} 24xy \, dx \, dy = z^4.$$

Now differentiate,

$$f_Z(z) = F'(z) = 4z^3.$$