24400 HW6

Bin Yu

November 12, 2024

Question 1

(a)

$$\mu = E[X_i] = \int_0^1 x \cdot f(x)\, dx = \int_0^1 x \cdot 2x\, dx = 2\int_0^1 x^2\, dx = 2\left[\frac{x^3}{3}\right]_0^1 = 2\left(\frac{1}{3} - 0\right) = \frac{2}{3}.$$

$$E[X_i^2] = \int_0^1 x^2 \cdot f(x)\, dx = \int_0^1 x^2 \cdot 2x\, dx = 2\int_0^1 x^3\, dx = 2\left[\frac{x^4}{4}\right]_0^1 = \frac{1}{2}.$$

$$\sigma^2 = E[X_i^2] - (E[X_i])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}.$$

For $n = 32$, the sum S is approximately normally distributed with:

$$n\mu = 32 \times \frac{2}{3} = \frac{64}{3}$$
$$n\sigma^2 = 32 \times \frac{1}{18} = \frac{16}{9}.$$

So, the normal distribution that approximates S is:

$$S \sim N\left(\frac{64}{3}, \frac{16}{9}\right).$$

(b)

We want to approximate $P(S \leq 20)$.

$$F(x) = P(S \leq x) = P\left(\frac{S-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$P(S \leq 20) = \Phi\left(\frac{20 - \frac{64}{3}}{\frac{4}{3}}\right) = \Phi\left(\frac{-\frac{4}{3}}{\frac{4}{3}}\right) = \Phi(-1).$$

Using standard normal distribution tables:

$$\Phi(-1) = 1 - \Phi(1) \approx 0.1587.$$

Therefore,

$$P(S \leq 20) \approx 0.1587.$$

Question 2

(a)

$$\mu_Y = E\left[\frac{X}{60}\right] = \frac{E[X]}{60} = \frac{np}{60} = \frac{60 \times 0.2}{60} = 0.2$$

$$\mathrm{Var}(Y) = \frac{p(1-p)}{n} = \frac{\frac{1}{5} \times \frac{4}{5}}{60} = \frac{\frac{4}{25}}{60} = \frac{1}{375}.$$

Therefore, according to the CLT, Y is approximately normally distributed:

$$Y \sim N\left(\mu_Y, \sigma_Y^2\right) = N\left(0.2, \frac{1}{375}\right).$$

(b)

To find$P\left(Y \leq 0.25\right)$:

$$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{\frac{1}{375}} = \frac{1}{\sqrt{375}} = \frac{1}{5\sqrt{15}}.$$

$$Z = \frac{Y - \mu_Y}{\sigma_Y} = \frac{0.25 - 0.2}{\frac{1}{5\sqrt{15}}} = \frac{5\sqrt{15}}{20} = \frac{\sqrt{15}}{4}.$$

$$Z = \frac{\sqrt{15}}{4} \approx 0.968.$$

$$P\left(Y \leq 0.25\right) = P\left(Z \leq \frac{\sqrt{15}}{4}\right) \approx P(Z \leq 0.968) \approx 0.8336.$$

Question 3

(a)

Let X_i be the net earnings from the i-th game. Then X_i is a random variable with the following distribution:

$$X_i = \begin{cases} 3 & \text{with probability } \frac{1}{6}, \\ -1 & \text{with probability } \frac{5}{6}. \end{cases}$$

Play the game $n = 36$ times, so the total earnings are:

$$S = \sum_{i=1}^{36} X_i.$$

We want to find $P(S > 0)$.

$$\mu_X = E[X_i] = \left(3 \times \frac{1}{6}\right) + \left(-1 \times \frac{5}{6}\right) = -\frac{1}{3}.$$

$$E[X_i^2] = \left(3^2 \times \frac{1}{6}\right) + \left((-1)^2 \times \frac{5}{6}\right) = \left(9 \times \frac{1}{6}\right) + \left(1 \times \frac{5}{6}\right) = \frac{9}{6} + \frac{5}{6} = \frac{7}{3}.$$

$$\sigma_X^2 = E[X_i^2] - (E[X_i])^2 = \frac{7}{3} - \left(-\frac{1}{3}\right)^2 = \frac{7}{3} - \frac{1}{9} = \frac{21}{9} - \frac{1}{9} = \frac{20}{9}.$$

Since the X_i are independent and identically distributed, the mean and variance of S are:

$$\mu_S = n\mu_X = 36 \times \left(-\frac{1}{3}\right) = -12.$$

$$\sigma_S^2 = n\sigma_X^2 = 36 \times \frac{20}{9} = 80.$$

$$\sigma_S = \sqrt{\sigma_S^2} = \sqrt{80} = 4\sqrt{5}.$$

According to the CLT, for large n, S is approximately normally distributed:

$$S \sim N\left(\mu_S, \sigma_S^2\right) = N\left(-12, 80\right).$$

$$P(S > 0) = P\left(\frac{S - \mu_S}{\sigma_S} > \frac{0 - (-12)}{4\sqrt{5}}\right).$$

$$Z = \frac{S - \mu_S}{\sigma_S} = \frac{0 - (-12)}{4\sqrt{5}} = \frac{12}{8.944} \approx 1.3416.$$

$$P(Z > 1.3416) = 1 - P(Z \leq 1.3416) = 1 - \Phi(1.3416)$$

Therefore:

$$P(Z > 1.3416) = 1 - 0.9102 = 0.0898.$$

The approximate probability that you are ahead after 36 games is about 8.98%.

(b)

In each round of the game:

$$W_i = 0.1(F_i - 20) = 0.1F_i - 2.$$

We play the game $n = 50$ times, so the total earnings are:

$$S = \sum_{i=1}^{50} W_i.$$

We want to find:

$$P(S \geq -30).$$

For an Exponential distribution with rate $\lambda = 0.1$:

$$E[F_i] = \frac{1}{\lambda} = \frac{1}{0.1} = 10.$$

$$\text{Var}(F_i) = \frac{1}{\lambda^2} = \frac{1}{(0.1)^2} = 100.$$

$$\mu_W = E[W_i] = E[0.1F_i - 2] = 0.1E[F_i] - 2 = 0.1 \times 10 - 2 = 1 - 2 = -1.$$

$$\sigma_W^2 = \text{Var}(W_i) = \text{Var}(0.1F_i - 2) = (0.1)^2 \,\text{Var}(F_i) = 0.01 \times 100 = 1.$$

Since the W_i are independent and identically distributed, the mean and variance of S are:

$$\mu_S = n\mu_W = 50 \times (-1) = -50.$$

$$\sigma_S^2 = n\sigma_W^2 = 50 \times 1 = 50.$$

$$\sigma_S = \sqrt{\sigma_S^2} = \sqrt{50}$$

According to the CLT, for large n, S is approximately normally distributed:

$$S \sim N\left(\mu_S, \sigma_S^2\right) = N\left(-50, 50\right).$$

$$P(S \geq -30) = P\left(\frac{S-\mu_S}{\sigma_S} \geq \frac{-30-(-50)}{\sqrt{50}}\right).$$

$$Z = \frac{S-\mu_S}{\sigma_S} = \frac{-30-(-50)}{\sqrt{50}} = \frac{20}{7.0711} \approx 2.8284.$$

$$P(Z \geq 2.8284) = 1 - \Phi(2.8284) = 1 - 0.9976 = 0.0024.$$

The approximate probability losing no more than \$30 after 50 games is approximately 0.24%.

(c)

We need to find:

$$P(S_{\text{total}} > -70)$$

Let:

- S_{game1} be the total earnings from the first game.
- S_{game2} be the total earnings from the second game.
- $S_{\text{total}} = S_{\text{game1}} + S_{\text{game2}}$ be the combined total earnings.

From (a) and (b):

$$\begin{aligned} \text{Mean } \mu_{S_1} &= -12 \\ \text{Variance } \sigma_{S_1}^2 &= 80 \end{aligned}$$

$$\begin{aligned} \text{Mean } \mu_{S_2} &= -50 \\ \text{Variance } \sigma_{S_2}^2 &= 50 \end{aligned}$$

Since the two games are independent,

$$\mu_{\text{total}} = \mu_{S_1} + \mu_{S_2} = -12 + (-50) = -62$$

$$\sigma_{\text{total}}^2 = \sigma_{S_1}^2 + \sigma_{S_2}^2 = 80 + 50 = 130$$

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{total}}^2} = \sqrt{130}$$

$$S_{\text{total}} \sim N\left(-62, 130\right).$$

$$P\left(S_{\text{total}} > -70\right) = P\left(\frac{S_{\text{total}} - \mu_{\text{total}}}{\sigma_{\text{total}}} > \frac{-70-(-62)}{11.4018}\right)$$

$$Z = \frac{-70-(-62)}{11.4018} = \frac{-8}{11.4018} \approx -0.7016$$

$$P(Z > -0.7016) = \Phi(0.7016) \approx 0.7585$$

Therefore,

$$P(S_{\text{total}} > -70) \approx 75.85\%$$

Question 4

(a)

For a random variable X, the moment-generating function $M_X(t)$ is defined as:

$$M_X(t) = E[e^{tX}]$$

Since X is a Bernoulli random variable, it takes on two values:

$X = 1$ with probability p
$X = 0$ with probability $1 - p$

Therefore:

$$\begin{aligned} M_X(t) &= E[e^{tX}] \\ &= e^{t\times 1} \cdot P(X=1) + e^{t\times 0} \cdot P(X=0) \\ &= e^t \cdot p + e^0 \cdot (1-p) \\ &= pe^t + (1-p). \end{aligned}$$

Taking the first derivative:

$$\begin{aligned} M_X'(t) &= \frac{d}{dt}\left(pe^t + (1-p)\right) \\ &= pe^t. \end{aligned}$$

Set t to 0:

$$M_X'(0) = pe^0 = p.$$

$$E[X] = M_X'(0) = p.$$

Taking the second derivative:

$$\begin{aligned} M''_X(t) &= \frac{d}{dt}\left(M'_X(t)\right) = \frac{d}{dt}\left(pe^t\right) \\ &= pe^t. \end{aligned}$$

at $t = 0$:

$$M''_X(0) = pe^0 = p.$$

$$\text{Var}[X] = M''_X(0) - \left(M'_X(0)\right)^2 = p - p^2 = p(1-p).$$

Taking the third derivative:

$$\begin{aligned} M'''_X(t) &= \frac{d}{dt}\left(M''_X(t)\right) = \frac{d}{dt}\left(pe^t\right) \\ &= pe^t. \end{aligned}$$

at $t = 0$:

$$M'''_X(0) = pe^0 = p.$$

Therefore,

$$E[X^3] = M'''_X(0) = p.$$

(b)

Y is the sum of n independent Bernoulli random variables:

$$Y = X_1 + X_2 + \cdots + X_n.$$

$$\begin{aligned} M_Y(t) &= E\left[e^{tY}\right] \\ &= E\left[e^{tX_1} e^{tX_2} \dots e^{tX_n}\right] \\ &= \prod_{i=1}^{n} E\left[e^{tX_i}\right] \quad \text{(since the } X_i \text{ are independent)} \\ &= \left(M_X(t)\right)^n. \end{aligned}$$

For a Bernoulli random variable $X_i \sim \text{Bernoulli}(p)$, the moment-generating function is:

$$M_X(t) = E\left[e^{tX_i}\right] = pe^t + (1-p).$$

Therefore,

$$M_Y(t) = \left(pe^t + (1-p)\right)^n.$$

Taking the first derivative of $M_Y(t)$:

$$M_Y'(t) = \frac{d}{dt}M_Y(t) = \frac{d}{dt}\left(pe^t + (1-p)\right)^n.$$

Let $u = pe^t + (1-p)$, so $M_Y(t) = u^n$.

Then,

$$\begin{aligned} M_Y'(t) &= nu^{n-1} \cdot \frac{du}{dt} \\ &= n\left(pe^t + (1-p)\right)^{n-1} \cdot \left(pe^t\right). \end{aligned}$$

at $t = 0$:

$$\begin{aligned} M_Y'(0) &= n\left(pe^0 + (1-p)\right)^{n-1} \cdot \left(pe^0\right) \\ &= n\left(p + 1 - p\right)^{n-1} \cdot p \\ &= np. \end{aligned}$$

$$E[Y] = M_Y'(0) = np.$$

Taking the second derivative of $M_Y(t)$:

$$M_Y''(t) = \frac{d^2}{dt^2}M_Y(t) = \frac{d}{dt}M_Y'(t).$$

$$M_Y'(t) = n\left(pe^t + (1-p)\right)^{n-1} \cdot pe^t.$$

Let $u = pe^t + (1-p)$ and $v = pe^t$, so $M_Y'(t) = nu^{n-1}v$.

$$\begin{aligned} M_Y''(t) &= n\left[\frac{d}{dt}\left(u^{n-1}\right) \cdot v + u^{n-1} \cdot \frac{dv}{dt}\right] \\ &= n\left[(n-1)u^{n-2} \cdot \frac{du}{dt} \cdot v + u^{n-1} \cdot \frac{dv}{dt}\right]. \\ &= n\left[(n-1)u^{n-2} \cdot pe^t \cdot v + u^{n-1} \cdot pe^t\right] \\ &= n\left[(n-1)\left(pe^t + (1-p)\right)^{n-2} \cdot pe^t \cdot pe^t + \left(pe^t + (1-p)\right)^{n-1} \cdot pe^t\right]. \end{aligned}$$

at $t = 0$:

$$
\begin{aligned}
M_Y''(0) &= n\left[(n-1)\cdot 1\cdot (p)^2 + 1\cdot p\right] \\
&= n\left[(n-1)p^2 + p\right] \\
&= n\left(p + (n-1)p^2\right).
\end{aligned}
$$

$$
(M_Y'(0))^2 = (np)^2 = n^2p^2.
$$

$$
\begin{aligned}
\mathrm{Var}[Y] &= M_Y''(0) - (M_Y'(0))^2 \\
&= n\left(p + (n-1)p^2\right) - n^2p^2 \\
&= np + n(n-1)p^2 - n^2p^2 \\
&= np + n^2p^2 - np^2 - n^2p^2 \\
&= np - np^2 \\
&= np(1-p).
\end{aligned}
$$

Therefore, the variance of Y is:

$$
\mathrm{Var}[Y] = np(1-p).
$$

Question 5

(a)

Conditional on $N = n$, Y is the sum of n independent exponential random variables:**

$$
(Y \mid N = n) = \sum_{i=1}^{n} X_i.
$$

$$
\begin{aligned}
E\left[e^{tY} \mid N = n\right] &= E\left[e^{t\sum_{i=1}^{n} X_i} \mid N = n\right] \\
&= E\left[\prod_{i=1}^{n} e^{tX_i}\right] \\
&= \prod_{i=1}^{n} E\left[e^{tX_i}\right] \quad \text{(since the } X_i \text{ are independent)} \\
&= \left(E\left[e^{tX}\right]\right)^n.
\end{aligned}
$$

$$
\begin{aligned}
E\left[e^{tX}\right] &= \int_0^\infty e^{tx}\lambda e^{-\lambda x}\,dx \\
&= \lambda \int_0^\infty e^{(t-\lambda)x}\,dx \\
&= \lambda \left[\frac{e^{(t-\lambda)x}}{t-\lambda}\right]_0^\infty \qquad \text{(converges if } t<\lambda\text{)} \\
&= \lambda\left(0 - \frac{1}{t-\lambda}\right) \\
&= \frac{\lambda}{\lambda - t}.
\end{aligned}
$$

Therefore,

$$
E\left[e^{tY} \mid N=n\right] = \left(\frac{\lambda}{\lambda-t}\right)^n, \quad \text{for } t<\lambda.
$$

(b)

Using the law of iterated expectations (Tower Law):

$$
E[e^{tY}] = E\left[E\left[e^{tY} \mid N\right]\right] = \sum_{n=1}^{\infty} E\left[e^{tY} \mid N=n\right] P(N=n).
$$

Given that $N \sim \text{Geometric}(p)$ on $\{1,2,3,\dots\}$:

$$
P(N=n) = p(1-p)^{n-1}, \quad n = 1,2,3,\dots
$$

$$
\begin{aligned}
E[e^{tY}] &= \sum_{n=1}^{\infty}\left(\frac{\lambda}{\lambda-t}\right)^n \cdot p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty}\left(\frac{\lambda}{\lambda-t}\right)^n \cdot \frac{p}{1-p}\cdot(1-p)^n \\
&= \frac{p}{1-p}\sum_{n=1}^{\infty}\left(\frac{\lambda}{\lambda-t}(1-p)\right)^n.
\end{aligned}
$$

Let

$$
r = \frac{\lambda}{\lambda-t}(1-p).
$$

Since $t<\lambda$ and $0<p<1$, $0<r<1$.

$$
\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}.
$$

$$
\begin{aligned}
E[e^{tY}] &= \frac{p}{1-p} \cdot \frac{r}{1-r} \\
&= \frac{p}{1-p} \cdot \frac{\frac{\lambda}{\lambda - t}(1-p)}{1 - \frac{\lambda}{\lambda - t}(1-p)} \\
&= \frac{p}{1-p} \cdot \frac{\frac{\lambda(1-p)}{\lambda - t}}{\frac{(\lambda - t) - \lambda(1-p)}{\lambda - t}} \\
&= \frac{p}{1-p} \cdot \frac{\lambda(1-p)}{\lambda - t - \lambda(1-p)} \\
&= \frac{p}{1-p} \cdot \frac{\lambda(1-p)}{\lambda p - t}. \\
&= \frac{\lambda p}{\lambda p - t}.
\end{aligned}
$$

Thus, the moment-generating function of Y is:

$$
E[e^{tY}] = \frac{\lambda p}{\lambda p - t}, \quad \text{for } t < \lambda p.
$$

(c)

The MGF of Y is:

$$
M_Y(t) = E\left[e^{tY}\right] = \frac{\lambda p}{\lambda p - t}, \quad t < \lambda p.
$$

Since the MGF uniquely decide the pdf, and this is the MGF of an exponential distribution with rate parameter λp:

$$
M(t) = \frac{\theta}{\theta - t}, \quad t < \theta.
$$

Therefore, Y follows an exponential distribution with rate λp:

$$
Y \sim \text{Exponential}(\lambda p).
$$

The probability density function (pdf) of Y is:

$$
f_Y(y) = \lambda p \, e^{-\lambda p \, y}, \quad y \geq 0.
$$

Question 6

(a)

The cumulative distribution function (CDF) of the maximum of n i.i.d. random variables can be found by:

$$F_{U_{(n)}}(x) = [F_U(x)]^n \,,$$

Since $U_i \sim \text{Uniform}[0,1]$, the CDF of U_i is:

$$F_U(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } 0 \le x \le 1, \\ 1 & \text{if } x > 1. \end{cases}$$

Therefore, for $0 \le x \le 1$:

$$F_{U_{(n)}}(x) = (x)^n \,.$$

(b)

$$Z_n = n(1 - U_{(n)}).$$

Therefore,

$$P(Z_n \le z) = P\left(n(1 - U_{(n)}) \le z\right) = P\left(U_{(n)} \ge 1 - \frac{z}{n}\right).$$

Since $U_{(n)}$ has CDF $F_{U_{(n)}}(x)$, its survival function is:

$$P(U_{(n)} \ge y) = 1 - F_{U_{(n)}}(y) = 1 - y^n.$$

Therefore,

$$\begin{aligned} F_{Z_n}(z) &= P\left(U_{(n)} \ge 1 - \frac{z}{n}\right) \\ &= 1 - \left(1 - \frac{z}{n}\right)^n, \quad \text{for } 0 \le z \le n. \end{aligned}$$

(c)

To find:

$$\lim_{n\to\infty} F_{Z_n}(z) = \lim_{n\to\infty} \left[1 - \left(1 - \frac{z}{n}\right)^n\right].$$

we have this limit:

$$\lim_{n\to\infty} \left(1 - \frac{z}{n}\right)^n = e^{-z}.$$

Therefore,

$$\lim_{n\to\infty} F_{Z_n}(z) = 1 - e^{-z}.$$

The limiting cumulative distribution function of Z_n as $n \to \infty$ is:

$$F_Z(z) = 1 - e^{-z}, \quad \text{for } z \geq 0.$$

This is the CDF of the Exponential distribution with rate parameter $\lambda = 1$.

Therefore, as $n \to \infty$, Z_n converges in distribution to an Exponential(1) random variable.