MATH 3215 Assignment 4

1. The probability of error in the transmission of a binary digit over a communication channel is 1/1000. When transmitting a block of 1000 bits, the error of each bit occurs independently.

(a) What is the probability that at least 3 errors occur in total?

(b) What if Poisson approximation is used?

For each question, give an exact expression and an approximate decimal value (4 digits after the decimal point).

Binomial: $1 - 0.999^{1000} - 1000 \cdot 0.001 \cdot 0.999^{999} - \binom{1000}{2} \cdot 0.001^2 \cdot 0.999^{998} \approx 0.0802$.

Poisson: $1 - e^{-1} - e^{-1} - e^{-1}/2 \approx 0.0803$.

2. Let $f(i)$ be the PMF of the Poisson distribution with parameter $\lambda > 1$.

(a) Show that $f(i)$ first increases and then decreases as i increases.

(b) Which value(s) of i (in terms of λ) achieves the maximum of $f(i)$?

Recall that $f(i) = e^{-\lambda}\frac{\lambda^i}{i!}$, so $\frac{f(i+1)}{f(i)} = \frac{\lambda}{i+1}$. Therefore, $f(i) \geq f(i+1)$ if and only if $i \geq \lambda - 1$. In other words, $f(i)$ increases until i is the smallest integer that is at least $\lambda - 1$ and then decreases; this i achieves the maximum of $f(i)$. If λ is an integer, then both $\lambda - 1$ and λ achieve the maximum.

3. A type of light bulbs has a normally distributed output with mean 2,000 fc and standard deviation 80 fc.

(a) What is the probability that a bulb has an output between 1,950 fc and 2,030 fc?

(b) What lower specification limit L fc can we set so that 95 percent of the bulbs produced will be above this limit?

For each question, give an exact expression in terms of the CDF Φ of $\mathcal{N}(0,1)$ and an approximate decimal value (2 digits after the decimal point).

Let the output be denoted by $X \sim \mathcal{N}(2000, 80^2)$. Then $Z := \frac{X-2000}{80} \sim \mathcal{N}(0,1)$. We have

$$\mathbb{P}\{1950 \leq X \leq 2030\} = \mathbb{P}\left\{\frac{-50}{80} \leq Z \leq \frac{30}{80}\right\} = \Phi(3/8) - \Phi(-5/8) \approx 0.38.$$

Moreover, we need

$$\mathbb{P}\{X \geq L\} = \mathbb{P}\left\{Z \geq \frac{L-2000}{80}\right\} = 1 - \Phi\left(\frac{L-2000}{80}\right) = 0.95.$$

Therefore,

$$L = 80 \cdot \Phi^{-1}(0.05) + 2000 \approx 1868.41.$$

4. What is $\mathbb{P}\{10000 < X < 20000\}$ if $X = e^Y$ where $Y \sim \mathcal{N}(10,1)$?

Give an exact expression in terms of the CDF Φ of $\mathcal{N}(0,1)$ and an approximate decimal value (3 digits after the decimal point).

$$\begin{aligned}
\mathbb{P}\{10000 < X < 20000\} &= \mathbb{P}\{10000 < e^Y < 20000\} \\
&= \mathbb{P}\{(\log 10000) - 10 < Y - 10 < (\log 20000) - 10\} \\
&= \Phi((\log 20000) - 10) - \Phi((\log 10000) - 10) \\
&\approx 0.247
\end{aligned}$$

5. Scores of a test across a large population follows the normal distribution with mean 500 and standard deviation 100. Five people are chosen randomly from the population.

(a) What is the probability that exactly three people scored above 580?

(b) What is the probability that at least four people scored below 600?

For each question, give an exact expression in terms of the CDF Φ of $\mathcal{N}(0,1)$ and an approximate decimal value (3 digits after the decimal point). (Hint: This problem involves both the normal and the binomial distribution.)

Let a score be denoted by $X \sim \mathcal{N}(500, 100^2)$. Then $Z := \frac{X-500}{100} \sim \mathcal{N}(0,1)$. We have

$$\mathbb{P}\{X > 580\} = \mathbb{P}\left\{Z > \frac{80}{100}\right\} = 1 - \Phi(0.8).$$

The number of people Y_1 who scored above 580 follows $\mathsf{Bin}\big(5, 1-\Phi(0.8)\big)$. Therefore,

$$\mathbb{P}\{Y_1 = 3\} = \binom{5}{3}\big(1-\Phi(0.8)\big)^3\,\Phi(0.8)^2 \approx 0.059.$$

Similarly,

$$\mathbb{P}\{X < 610\} = \mathbb{P}\left\{Z < \frac{100}{100}\right\} = \Phi(1).$$

The number of people Y_2 who scored below 600 follows $\mathsf{Bin}\big(5, \Phi(1)\big)$. Therefore,

$$\mathbb{P}\{Y_2 \geq 4\} = \binom{5}{4}\Phi(1)^4\big(1-\Phi(0.8)\big) + \Phi(1)^5 \approx 0.819.$$

6. For $p \in (0,1)$, suppose that we observe a sequence of independent $\mathsf{Ber}(p)$ trials (i.e., each trial succeeds with probability p) until the first success occurs. Let X denote the total number of trials. For example, if the first two trials fail and the third succeeds, then $X = 3$. Determine the sample space, PMF, CDF, MGF ($M(t)$ for $t < \log\frac{1}{1-p}$), mean, and variance of X.

(Recall the formula for a geometric sum $\sum_{k=1}^{n} r^{k-1} = \frac{1-r^n}{1-r}$ for $r \neq 1$. When computing $M'(t)$ and $M''(t)$, it is okay not to show intermediate steps.)

The sample space is the set of positive integers.

The PMF is $f(i) = (1-p)^{i-1}p$ for $i = 1, 2, 3, \ldots$

The CDF is $F(i) = \sum_{j=1}^{i}(1-p)^{j-1}p = \frac{1-(1-p)^i}{1-(1-p)}p = 1-(1-p)^i$ for $i = 1, 2, 3, \ldots$ To be more precise, for $i \leq x < i+1$, we have $F(x) = F(i)$, and for $x < 1$, we have $F(x) = 0$.

The MGF is $M(t) = \sum_{i=1}^{\infty} e^{ti}(1-p)^{i-1}p = e^t p \sum_{i=1}^{\infty} \left(e^t(1-p)\right)^{i-1} = \frac{e^t p}{1-e^t(1-p)}$.

The mean is $\mathbb{E}[X] = M'(0) = \frac{e^t p}{(1-e^t(1-p))^2}\Big|_{t=0} = 1/p$.

The second moment is $\mathbb{E}[X^2] = M''(0) = \frac{e^t p(1+e^t(1-p))}{(1-e^t(1-p))^3}\Big|_{t=0} = \frac{2-p}{p^2}$.

The variance is $\mathsf{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1-p}{p^2}$.

7. Suppose that the joint PDF of (X, Y) is $f(x, y) = c\big(x^2 + \frac{xy}{2}\big)$ for $x \in (0, 1)$, $y \in (0, 2)$, and $f(x, y) = 0$ otherwise, where c is a constant.

(a) What is the constant c?

(b) Find $\mathbb{P}\{X > Y\}$.

We have

$$1 = \iint f(x, y)\, dx\, dy = \int_0^2 \int_0^1 c\Big(x^2 + \frac{xy}{2}\Big)\, dx\, dy = c\Big(\frac{x^3 y}{3} + \frac{x^2 y^2}{8}\Big)\Big|_{x=0}^{1}\Big|_{y=0}^{2} = c\,\frac{7}{6},$$

so $c = 6/7$. Moreover,

$$\mathbb{P}\{X > Y\} = \int_0^1 \int_y^1 \frac{6}{7}\Big(x^2 + \frac{xy}{2}\Big)\, dx\, dy = \int_0^2 \frac{-7y^3 + 3y + 4}{14}\, dy = \frac{15}{56}.$$

8. Let (X, Y) have joint PDF $f(x, y) = xe^{-x-y}$ for $x > 0$, $y > 0$ and $f(x, y) = 0$ otherwise.

(a) Compute the marginal PDFs of X and Y.

(b) Determine whether X and Y are independent.

We have

$$f_X(x) = \int_0^\infty xe^{-x-y}\, dy = xe^{-x}, \qquad f_Y(y) = \int_0^\infty xe^{-x-y}\, dx = e^{-y},$$

for $x > 0$ and $y > 0$ respectively. Hence $f(x, y) = f_X(x) \cdot f_Y(y)$, so X and Y are independent.