CHEM. ENG. II : PROBABILITY AND STATISTICS

Chapter 3. Special Discrete Probability Distributions

Discrete probability models are used to model the outcomes of counting experiments. Depending on the experimental situation, it is often possible to justify the use of one of a class of "Special" discrete probability distributions. These are listed in this chapter, and are all motivated from the central concept of a *binary* or 0-1 trial, where the random variable concerned has range consisting of only two values with associated probabilities θ and $1-\theta$ respectively; typically we think of the possible outcomes as "successes" and "failures".

All of the distributions in this chapter are derived by making different modelling assumptions about sequences of 0-1 trials.

Single 0-1 trial - count number of 1s	$\Longrightarrow$ BERNOULLI DISTRIBUTION
n independent 0-1 trials - count number of 1s	$\Longrightarrow$ BINOMIAL DISTRIBUTION
Sequence of independent 0-1 trials - count number of trials until first 1	$\Longrightarrow$ GEOMETRIC DISTRIBUTION
Sequence of independent 0-1 trials - count number of trials until nth 1	$\Longrightarrow$ NEGATIVE BINOMIAL DISTRIBUTION
Limiting case of binomial distribution	$\Longrightarrow$ POISSON DISTRIBUTION

3.1 The Bernoulli Distribution $X \sim Bernoulli(\theta)$

Range : $\mathbb{X} = \{0,1\}$
Parameter : $\theta \in [0,1]$
Mass function :

$$f_X(x) = \theta^x(1-\theta)^{1-x} \qquad x \in \{0,1\}$$

Interpretation : A model for "the number of successes in a single 0-1 trial" where the probability of success is θ.

EXPECTATION $\quad \mathrm{E}_{f_X}[\ X\ ] = \sum_{x=0}^{1} x\theta^x(1-\theta)^{1-x} = \theta$

VARIANCE $\quad \mathrm{Var}_{f_X}[\ X\ ] = \mathrm{E}_{f_X}[\ X^2\ ] - \{\mathrm{E}_{f_X}[\ X\ ]\}^2 = \sum_{x=0}^{1} x^2\theta^x(1-\theta)^{1-x} - \theta^2 = \theta(1-\theta)$

3.2 The Binomial Distribution $X \sim Binomial(n,\theta)$

Range : $\mathbb{X} = \{0,1,...,n\}$
Parameters : $n \in Z^+$, $\theta \in [0,1]$
Mass function :

$$f_X(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x} = \frac{n!}{x!(n-x)!}\theta^x(1-\theta)^{n-x} \qquad x \in \{0,1,...,n\}$$

Recall that $f_X(x) = \mathrm{P}[\ X = x\ ]$. Now if $X = x$, we must have a sequence of n 0-1 trials which contains x 1s and $n - x$ 0s. As the trials are independent, with probability θ of obtaining a 1 on any trial, all such sequences have probability

$$\theta^x(1-\theta)^{n-x}$$

and there are ${}^nC_x = \binom{n}{x}$ such sequences; we must choose the x positions in the sequence to place the ones. Hence the total probability of observing the event $X = x$ is

$$\binom{n}{x}\theta^x(1-\theta)^{n-x}$$

Interpretation : A model for "the number of successes in n i.i.d. 0-1 trials" where the probability of success on each trial is θ. Note that the sum of n independent and identically distributed (i.i.d.) $Bernoulli(\theta)$ random variables has a Binomial distribution, that is,

$$\text{If } X_1, ..., X_n \sim Bernoulli(\theta)\ i.i.d., \text{ then } X = \sum_{i=1}^{n} X_i \sim Binomial(n, \theta)$$

It is quite straightforward to calculate the expectation and variance direct from the definitions, but easier to use the results for sums of independent random variables from the previous chapter

$$\text{EXPECTATION} \quad \mathrm{E}_{f_X}[\ X\ ] = \sum_{x=1}^{n} \mathrm{E}_{f_{X_i}}[\ X_i\ ] = \sum_{x=1}^{n} \theta = n\theta$$

$$\text{VARIANCE} \quad \mathrm{Var}_{f_X}[\ X\ ] = \sum_{x=1}^{n} \mathrm{Var}_{f_{X_i}}[\ X_i\ ] = \sum_{x=1}^{n} \theta(1-\theta) = n\theta(1-\theta)$$

EXAMPLE : In a multiple choice exam, each of the ten questions has five possible answers. For any single question, the probability of guessing the right answer is $1/5$. Therefore the random variable X corresponding to the total number of correct answers assuming that each question is answered by guesswork has a binomial distribution, $X \sim Binomial(10, 1/5)$.

3.3 The Geometric Distribution $X \sim Geometric(\theta)$

Range : $\mathbb{X} = \{1, 2, ...\}$
Parameter : $\theta \in (0, 1]$
Mass function :

$$f_X(x) = (1-\theta)^{x-1}\theta. \qquad x \in \{1, 2, ...\}$$

Interpretation : A model for "the number of 0-1 trials until the first success is obtained " - where the probability of success on each trial is θ. a **discrete waiting-time** model.

Recall that $f_X(x) = \mathrm{P}[\ X = x\ ]$. Now if $X = x$, we must have a sequence of x 0-1 trials which contains $x - 1$ 0s and then a 1. As the trials are independent, with probability θ of obtaining a 1 on any trial, such a sequence has probability

$$(1-\theta)^{x-1}\theta$$

which immediately gives the form of the mass function.

EXAMPLE : A coin is tossed until the first head is obtained. The random variable X corresponding to the number of tosses required to obtain a the first head has a geometric distribution, $X \sim Geometric(1/2)$

$$\text{EXPECTATION} \quad \mathrm{E}_{f_X}[\,X\,] = \sum_{x=1}^{\infty} x(1-\theta)^{x-1}\theta = \theta \sum_{x=1}^{\infty} \left\{ \frac{d}{d\theta} - (1-\theta)^x \right\} = -\theta \frac{d}{d\theta}\left\{ \sum_{x=0}^{\infty} (1-\theta)^x \right\}$$

$$= -\theta \frac{d}{d\theta}\left\{ \frac{1}{1-(1-\theta)} \right\}$$

$$= -\theta \, \frac{-1}{\theta^2} = \frac{1}{\theta}$$

$$\text{VARIANCE} \quad \mathrm{Var}_{f_X}[\,X\,] = \frac{1-\theta}{\theta^2}$$

3.4 <u>The Negative Binomial Distribution</u> $X \sim NegBin(n, \theta)$

Range : $\mathbb{X} = \{n, n+1, n+2, ...\}$
Parameter : $n \in Z^+$, $\theta \in (0, 1]$
Mass function :

$$f_X(x) = \left(\begin{array}{c} x-1 \\ n-1 \end{array} \right) \theta^n (1-\theta)^{x-n} \qquad x \in \{n, n+1, n+2, ...\}\,.$$

Interpretation : A model for "the number of 0-1 trials until the nth success is obtained" where the probability of success on each trial is θ.

To compute the mass function, note that if the nth success occurs on the xth trial, then the sequence of 0-1 trials must comprise a sequence of $x-1$ trials containing $n-1$ 1s and $x-n$ 0s, followed by a final 1. Thus the mass function for a negative binomial distribution is given by

$$\binom{x-1}{n-1} \theta^{n-1}(1-\theta)^{x-n} \times \theta$$

that is a binomial probability multiplied by a single θ. as all trials are independent.

Note: The sum of n i.i.d. $Geometric(\theta)$ random variables has a Negative Binomial distribution, that is,

$$\text{If } X_1, ..., X_n \sim Geometric(\theta) \text{ } i.i.d., \text{ then } X = \sum_{i=1}^{n} X_i \sim NegBin(n, \theta)$$

that is, the number of trials until the nth 1 is the sum of the number of trials until the first 1, plus the number of trials between the first and second 1, etc.

EXAMPLE : A coin is tossed until the tenth head is obtained. The random variable X corresponding to the number of tosses required to obtain a the tenth head has a negative binomial distribution, $X \sim NegBin(10, 1/2)$

It is quite straightforward to calculate the expectation and variance direct from the definitions, but easier to use the results for sums of independent random variables from the previous chapter

$$\text{EXPECTATION} \quad \mathrm{E}_{f_X}[\,X\,] = \sum_{x=1}^{n} \mathrm{E}_{f_{X_i}}[\,X_i\,] = \sum_{x=1}^{n} \frac{1}{\theta} = \frac{n}{\theta}$$

$$\text{VARIANCE} \quad \mathrm{Var}_{f_X}[\,X\,] = \sum_{x=1}^{n} \mathrm{Var}_{f_{X_i}}[\,X_i\,] = \sum_{x=1}^{n} \frac{(1-\theta)}{\theta^2} = \frac{n(1-\theta)}{\theta^2}$$

3.5 <u>The Poisson Distribution</u> $X \sim Poisson(\lambda)$

Range : $\mathbb{X} = \{0, 1, 2, ...\}$
Parameter : $\lambda \in \mathbb{R}^+$
Mass function :

$$f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad x \in \{0, 1, 2, ...\}$$

Note : The Poisson is the limiting case of the binomial distribution, that is, if we write $\theta = \lambda/n$, and then consider a limiting case as $n \longrightarrow \infty$. Specifically, if $X \sim Bin(n, \lambda/n)$, then

$$\begin{aligned} f_X(x) &= \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n!}{(n-x)!} \left(\frac{1}{n-\lambda}\right)^x \\ &\longrightarrow \frac{\lambda^x}{x!} e^{-\lambda} \end{aligned}$$

as $n \longrightarrow \infty$ with λ constant. That is, we have that

$$Binomial(n, \theta) \longrightarrow Poisson(\lambda).$$

in the limiting case.

Interpretation : The Poisson model is appropriate for count data, where the number of events (accidents, breakdowns etc) that occur in a given time period are being counted.

Alternative derivation : Consider a sequence of events occurring **independently** and **at random** in time at a **constant rate** λ per unit time. Let $X(t)$ be the random variable defined for $t > 0$ by

"$X(t) = x$" if and only if "x events occur in time interval $[0, t)$".

Then $X(t) \sim Poisson(\lambda t)$, so that

$$f_{X(t)}(x) = \mathrm{P}[X(t) = x] = \frac{(\lambda t)^x}{x!} e^{-(\lambda t)} \qquad x \in \{0, 1, 2, ...\}.$$

EXAMPLE : The reliability of a machine is to be monitored. The random variable X corresponding to the number of breakdowns that the machine suffers in one month can be modelled using a Poisson distribution.

The expectation and variance calculations are quite straightforward. To calculate the expectation, we use the power series expansion for the exponential function;

$$\text{EXPECTATION } \mathrm{E}_{f_X}[\ X\ ] = \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

For the variance, we first note that

$$\mathrm{Var}_{f_X}[\ X\ ] = \mathrm{E}_{f_X}[\ (X - \{\mathrm{E}_{f_X}[\ X\ ]\})^2\ ] = \mathrm{E}_{f_X}[\ X^2\ ] - \{\mathrm{E}_{f_X}[\ X\ ]\}^2 = \mathrm{E}_{f_X}[\ X(X-1) + X\ ] - \{\mathrm{E}_{f_X}[\ X\ ]\}^2$$

and hence, using the properties of expectation

$$\mathrm{Var}_{f_X}[\ X\ ] = \mathrm{E}_{f_X}[\ X(X-1)\ ] + \mathrm{E}_{f_X}[\ X\ ] - \lambda^2$$

where

$$\mathrm{E}_{f_X}[\ X(X-1)\ ] = \sum_{x=2}^{\infty} x(x-1) \frac{\lambda^x}{x!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2$$

so

$$\text{VARIANCE } \mathrm{Var}_{f_X}[\ X\ ] = \lambda^2 + \lambda - \lambda^2 = \lambda$$