MATH 353

PROBABILITY AND STATISTICS

UNIT 4

Alexander Boateng, PhD

Department of Statistics and Actuarial Science, KNUST

UNIT 4: OUTLINE

1. Binomial Probability Distribution
2. Poisson Probability Distribution
3. Discrete Uniform Probability
4. Geometric Probability Distribution
5. Normal Probability Distribution

UNIT OBJECTIVES:

After completing this unit, you should be able to:

- Explain and compute **binomial probabilities** using the binomial model.
- Explain and compute **Poisson probabilities** using the Poisson model.
- Explain and compute normal probabilities using the **standard normal probability table**.
- Calculate the **mean, variance and standard deviation** of these distributions..

BINOMIAL PROBABILITY DISTRIBUTION

Definition:

- This is used to model experiments consisting of a sequence of observations of ***identical and independent trials***, each of which results in one of the possible outcomes.

Conditions for a Binomial Model

- A finite number, n, trials are carried out.
- The trials are independent.
- The outcome of each trial is deemed either a success or failure.

CONDITIONS FOR A BINOMIAL MODEL cont'd.

- The probability, p, of a successful outcome is the same for each trial.
- The discrete random variable, X, is **the number of successful outcomes in n trials.**
- If these conditions are satisfied, X is said to follow a **binomial distribution** denoted as:

$$X \sim B(n, p)$$

DISTRIBUTION FOR A BINOMIAL MODEL

- The distribution of the binomial model is:

$$P(X = x) = nC_x\, p^x\, (1 - p)^{n-x}$$

$$\text{for } x = 0, 1, 2, 3, \ldots, n$$

- **NB:** $\boldsymbol{p}$ is the probability of success

  $\boldsymbol{n}$ is the number of trials

  $\boldsymbol{q = 1 - p}$ is the probability of failure

  $\boldsymbol{nC_x}$ is same as $\binom{n}{x}$ i.e. $\boldsymbol{n}$ **Combination** $\boldsymbol{x}$

  $\boldsymbol{n}$ and $\boldsymbol{p}$ are the parameters for binomial probability.

Example 4.1

- Given that the random variable X is binomially distributed with $n = 10$ and $p = 0.2$. That is $X \sim B(10, 0.2)$.

a) Find $P(X = 0)$.

b) Find $P(X \leq 3)$.

c) Find $P(X > 3)$.

d) Find $P(2 < X \leq 5)$.

e) Find $P(X = 10)$.

Solution 4.1

$\boldsymbol{X \sim B(10, 0.2)}$ **where** $\boldsymbol{n = 10}$ **and** $\boldsymbol{p = 0.2 \quad => q = 1 - p = 1 - 0.2 = 0.8}$

a) $P(X = 0) = 10C0\ 0.2^0 (1 - 0.2)^{10-0} = 1 \times 1 \times 0.8^{10} = 0.1074$

b) $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

$$= 10C0\ 0.2^0 (1 - 0.2)^{10-0} + 10C1\ 0.2^1 (1 - 0.2)^{10-1} + 10C2\ 0.2^2 (1 - 0.2)^{10-2} + 10C3\ 0.2^3 (1 - 0.2)^{10-3}$$

$$= 0.8791$$

OR compactly as:

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$= \sum_{x=0}^{3} 10Cx\ 0.2^x (1 - 0.2)^{10-x}$$

$$= 0.8791$$

Solution 4.1 cont'd.

c) $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.8791 = 0.1209$

OR one can compute for $P(X \geq 4)$

$$= \sum_{x=4}^{10} 10Cx\ 0.2^{x}(1-0.2)^{10-x} = 0.1209$$

d) $P(2 < X \leq 5) = P(3 \leq X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5)$

$$= 10C0\ 0.2^{3}(1-0.2)^{10-3} + 10C1\ 0.2^{4}(1-0.2)^{10-4}$$
$$+ 10C2\ 0.2^{5}(1-0.2)^{10-5}$$
$$= 0.3158$$

Solution 4.1 cont'd.

OR compactly as:

d) $P(2 < X \leq 5) = P(3 \leq X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5)$

$$= \sum_{x=3}^{5} 10Cx\ 0.2^{x}(1-0.2)^{10-x}$$

$$= 0.3158$$

e) $P(X = 10) = 10C10\ 0.2^{10}(1-0.2)^{10-10} = 1 \times 0.2^{10} \times 1 = 0.2^{10}$

$$= 1.02 \times 10^{-7}$$

Example 4.2

30% of students in college of engineering in a particular higher institution travels to the central class by bus provided by the institute. To estimate the probability of a certain fraction of a class in the college traveling to class by bus, a lecturer chose fifteen students at random from this class in the college.

a) Find the probability that only three students travel to class by bus.
b) Find the probability that less than five students travel to class by bus.
c) Find the probability that not fewer than ten students travel to class by bus.
d) Find the probability that at least ten students travel to class by bus.
e) Find the probability that between six and ten students travel to class by bus.
f) Find the probability that at most five students travel to class by bus.

Solution 4.2

Let X be the number of students who travel to class by bus in the class.

$p = 30\% = 0.3$ and $q = 1 - p = 70\% = 0.7 \quad n = 15.$

$=> X \sim B(15, 0.3).$

a) $P(X = 3) = 15C3\ 0.3^3(1 - 0.3)^{15-3} = 0.1700$

b) $P(X < 5) = P(X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= 15C0\ 0.3^0(1 - 0.3)^{15-0} + 15C1\ 0.3^1(1 - 0.3)^{15-1} + 15C2\ 0.3^2(1 - 0.3)^{15-2} + 15C3\ 0.3^3(1 - 0.3)^{15-3} + 15C4\ 0.3^4(1 - 0.3)^{15-4}$$

$$= 0.5155$$

Solution 4.2 cont'd.

c) $P(X \geq 10) = P(X = 10) + P(X = 11) + P(X = 12) + P(X = 13) + P(X = 14) + P(X = 15)$

$$= \sum_{x=10}^{15} 15Cx\, 0.3^x (1 - 0.3)^{15-x}$$

$$= 0.0037$$

d) $P(X \geq 10) = P(X = 10) + P(X = 11) + P(X = 12) + P(X = 13) + P(X = 14) + P(X = 15)$

$$= \sum_{x=10}^{15} 15Cx\, 0.3^x (1 - 0.3)^{15-x}$$

$$= 0.0037$$

Solution 4.2 cont'd.

e) $P(6 < X < 10) = P(X = 7) + P(X = 8) + P(X = 9)$

$$= 15C7\, 0.3^7(1-0.3)^{15-7} + 15C8\, 0.3^8(1-0.3)^{15-8}$$
$$+ 15C9\, 0.3^9(1-0.3)^{15-9}$$
$$= 0.1275$$

f) $P(X \leq 5) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) +$
$P(X = 4) + P(X = 5)$

$$= 15C0\, 0.3^0(1-0.3)^{15-0} + 15C1\, 0.3^1(1-0.3)^{15-1}$$
$$+ 15C2\, 0.3^2(1-0.3)^{15-2} + 5C3\, 0.3^3(1-0.3)^{15-3}$$
$$+ 15C4\, 0.3^4(1-0.3)^{15-4} + 15C5\, 0.3^5(1-0.3)^{15-5}$$
$$= 0.7216$$

EXPECTATION AND VARIANCE OF BINOMIAL DISTRIBUTION

The expectation (mean) and the variance of the binomial distribution $(X \sim B(n, p))$ are defined as:

- $\text{Expectation} = E(X) = np$
- $\text{Varaince} = \sigma^2 = Var(X) = np(1-p) = npq$ where $q = 1 - p$
- $Standard\ deviation = \sigma = SD(X) = \sqrt{\text{Variance}} = \sqrt{npq}$

Example 4.3

- 10% of the articles from a certain production line are defective. A sample of 25 articles is taken.
  a) Find the expected number of defective articles.
  b) Find the variance of the number of defective articles.
  c) Find the standard deviation of the number of the defective articles.

Solution 4.3

- Let X be the number of defective articles.
- => $X \sim B(25, 0.1)$
- $E(X) = np = 25 \times 0.1 = 2.5$
- $Var(X) = npq = 25 \times 0.1 \times (1 - 0.1) = 25 \times 0.1 \times 0.9 = 2.25$
- $SD(X) = \sqrt{Var(X)} = \sqrt{2.25} = 1.5$

- NB: Sometimes, the mean, variance or standard deviation and probability of the binomial model, one or two of these will be given and you will be asked to find p and n.

Example 4.4

- The random variable X is $B(n, 0.6)$ and $P(X < 1) = 0.0256$.
- Find the value of n.

Solution 4.4

- $P(X < 1) = P(X = 0) = nC0\ 0.6^0(1 - 0.6)^{n-0} = 0.0256$
- $= 1 \times 1 \times 0.4^n = 0.0256$
- $= \ln 0.4^n = \ln 0.0256$
- $n = \dfrac{\ln 0.0256}{\ln 0.4} = 4$
- $\therefore n = 4\ and\ X{\sim}B(4, 0.6)$

Example 4.5

- Each day a bakery delivers the same number of loaves to a certain shop which sells 98% of them. Assuming that the number of loaves sold per day has a binomial distribution with a standard deviation of 7.

a) Find the number of loaves the shop would expect to sell per day.

b) Find the probability that only 1% of the number of loaves are sold per day.

Solution 4.5

- Let X be the number of loaves sold.
- $=> X \sim B(n, 0.98)$
- $SD(X) = 7 = \sqrt{npq} = \sqrt{n \times 0.98 \times 0.02}$
- $=> 0.0196n = 7^2$
- $n = \frac{49}{0.0196} = 2500$

a) The shop would expect to sell $np = 2500 \times 0.98 = 2450$

1% of 2500 = 25

b) $P(X = 25) = 2500C25\ (0.98)^{25}(1 - 0.98)^{2500-25} = 0$

THE POISSON DISTRIBUTION

- The Poisson distribution for a random variable, X, representing the number of occurrence of an event in a given interval of time, space or volume is defined by:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad for \;\; x = 0, 1, 2, 3, \ldots \infty$$

- Where, the mean and variance are equal.

$$E(X) = \lambda = Var(X)$$

NB:

- In some text, the mu, μ, is used in place of lambda, λ.
- The Poisson random variable takes values from **0 to infinity** unlike the Binomial model which takes values from **0 to n**.

PROBLEMS SUITABLE FOR THE POISSON R.V.

- The number of emergency calls received by an ambulance control in an hour.
- The number of vehicles approaching a motorway toll bridge in a five-minute interval.
- The number of flaws in a metre length of a material.
- The number of typed mistakes on a page of a book.

Conditions for a Poisson Model

- Events occur singly and at random in a given interval of time or space.
- λ, the mean number of occurrences in the given interval, is known and is finite.

Example 4.6

- Given that X follows a Poisson probability distribution with $\lambda = 2.5$.

a) Find $P(X = 10)$.

b) Find $P(X \leq 3)$.

c) Find $P(X > 3)$.

d) Find $P(X \geq 5)$.

e) Find $P(0 < X < 4)$.

Solution 4.6

$X \sim P_o(\lambda = 2.5)$.

a) $P(X = 10) = e^{-2.5}\frac{2.5^{10}}{10!} = 0.0002$

b) $P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

$$= e^{-2.5}\frac{2.5^0}{0!} + e^{-2.5}\frac{2.5^1}{1!} + e^{-2.5}\frac{2.5^2}{2!} + e^{-2.5}\frac{2.5^3}{3!}$$

$$= 0.7576$$

c) $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.7576 = 0.2424$

Solution 4.6 cont'd.

d) $P(X \geq 5) = 1 - P(X \leq 4)$

$$= 1 - \left[e^{-2.5}\frac{2.5^0}{0!} + e^{-2.5}\frac{2.5^1}{1!} + e^{-2.5}\frac{2.5^2}{2!} + e^{-2.5}\frac{2.5^3}{3!} + e^{-2.5}\frac{2.5^4}{4!}\right]$$

$$= 1 - 0.8912$$

$$= 0.1088$$

e) $P(0 < X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$

$$= e^{-2.5}\frac{2.5^1}{1!} + e^{-2.5}\frac{2.5^2}{2!} + e^{-2.5}\frac{2.5^3}{3!}$$

$$= 0.6755$$

Example 4.7

- On average the school photocopier breaks down eight times during the school weak (Monday to Friday). Assuming that the number of breakdowns can be modelled by a Poisson distribution.

  a) Find the probability that it breaks down five times in a given week.

  b) Find the probability that it breaks down not less than three times in a given week

  c) Find the probability that it breaks down once on Monday.

  d) Find the probability that it breaks down eight times in a fortnight.

Solution 4.7

Let X be the number of photocopier break down in a **week.**

=> $X \sim P_o(\lambda = 8)$.

a) $P(X = 5) = e^{-8}\frac{8^5}{5!} = 0.0916$

b) $P(X \geq 4) = 1 - P(X \leq 3)$

$$= 1 - \left[e^{-8}\frac{8^0}{0!} + e^{-8}\frac{8^1}{1!} + e^{-8}\frac{8^2}{2!} + e^{-8}\frac{8^3}{3!}\right]$$

$$= 1 - 0.0424$$

$$= 0.9576$$

Solution 4.7 cont'd.

c) Here X cannot model the question because X occurrence rate is weekly.

Let Y be the number of photocopier break down in a day.

$Y \sim P_o(\lambda = \frac{8}{5} = 1.6)$.

$$P(Y = 1) = e^{-1.6}\frac{1.6^1}{1!} = 0.3230$$

d) Let Z be the number of break downs in a fortnight.

$Z \sim P_o(\lambda = 2 \times 8 = 16)$.

$$P(Z = 8) = e^{-16}\frac{16^8}{8!} = 0.0120$$

EXPECTATION AND VARIANCE OF THE POISSON

- The expectation (mean) and the variance of the Poisson distribution ($X \sim P_o(\lambda)$) are defined as:

- $\text{Expectation} = E(X) = \lambda$
- $\text{Variance} = \sigma^2 = Var(X) = \lambda$
- $Standard\ deviation = \sigma = SD(X) = \sqrt{\text{Variance}} = \sqrt{\lambda}$

Example 4.8

- X follows a Poisson distribution with standard deviation 4.
  a) Find $P(X = 0)$.
  b) Find $P(X < 3)$.
  c) Find $P(X > 3)$.
  d) Find $P(X \geq 5)$.
  e) Find $P(0 < X < 4)$.

Solution 4.8

We need to find the Poisson rate, λ.

Variance = Mean = Poisson rate = (Standard deviation)^2

$\therefore \lambda = 4^2 = 16.$

$X \sim P_o(\lambda = 16).$

a) $P(X = 0) = e^{-16}\frac{16^0}{0!} = 1.13 \times 10^{-7}$

b) $P(X < 3) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$$= e^{-16}\frac{16^0}{0!} + e^{-16}\frac{16^1}{1!} + e^{-16}\frac{16^2}{2!}$$

$$= 1.63 \times 10^{-5}$$

Solution 4.8 cont'd.

- c) $P(X > 3) = 1 - P(X \leq 3)$

$$= 1 - \left[e^{-16}\frac{16^0}{0!} + e^{-16}\frac{16^1}{1!} + e^{-16}\frac{16^2}{2!} + e^{-16}\frac{16^3}{3!}\right]$$
$$= 1 - 9.31 \times 10^{-5}$$
$$= 0.9999$$

d) $P(X \geq 5) = 1 - P(X \leq 4)$

$$= 1 - \left[e^{-16}\frac{16^0}{0!} + e^{-16}\frac{16^1}{1!} + e^{-16}\frac{16^2}{2!} + e^{-16}\frac{16^3}{3!} + e^{-16}\frac{16^4}{4!}\right]$$
$$= 1 - 0.0004$$
$$= 0.9996$$

e) $P(0 < X \leq 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= e^{-16}\frac{16^1}{1!} + e^{-16}\frac{16^2}{2!} + e^{-16}\frac{16^3}{3!} + e^{-16}\frac{16^4}{4!}$$
$$= 0.0004$$

THE SUM OF INDEPENDENT POISSON R.V.

For independent Poisson variables X and Y,

i.e. if $X \sim P_o(\lambda_1)$ and $Y \sim P_o(\lambda_2)$,

Then:

- $X + Y \sim P_o(\lambda_1 + \lambda_2)$

Example 4.9

- Given that $X \sim P_o(4)$ and $Y \sim P_o(2.5)$.

  a) Find $P(X + Y = 3)$.

  b) Find $P(X - 0.5Y > 2)$. OR $P(X - \frac{1}{2}Y > 2)$

Solution 4.9

a) Distribution of $X+Y \sim P_o(4+2.5=6.5)$

$$P(X+Y=3)=e^{-6.5}\frac{6.5^3}{3!}=0.0688$$

a) Distribution of $X-0.5Y \sim P_o(4-0.5\times 2.5=2.75)$

$$P(X-0.5Y>2)=1-P(X-0.5Y\leq 2)$$

$$=1-\left[e^{-2.75}\frac{2.75^0}{0!}+e^{-2.75}\frac{2.75^1}{1!}+e^{-2.75}\frac{2.75^2}{2!}\right]$$

$$=1-0.4815$$

$$=0.5185$$

Example 4.10

- Telephone calls reach a secretary independently and at random, internal ones at a mean rate of two in any five-minute period, and external ones at a mean rate of one in any five-minute period.

  Calculate the probability that there will be more than two calls in any period of two minutes.

Solution 4.10

Let X and Y be internal and external number of calls reaching the secretary respectively.

=> $X \sim P_o(2)$ and $Y \sim P_o(1)$ for five-minute period.

∴ for two minutes period, $X \sim P_o(\frac{2\times 2}{5} = 0.8)$ and $Y \sim P_o(\frac{2}{5} = 0.4)$

Let T be total calls reaching the secretary.

$$=> X + Y = T$$

Hence $T \sim P_o(0.8 + 0.4 = 1.2)$ in two minutes

$$P(T > 2) = 1 - P(T \leq 2)$$

$$= 1 - \left[e^{-1.2}\frac{1.2^0}{0!} + e^{-1.2}\frac{1.2^1}{1!} + e^{-1.2}\frac{1.2^2}{2!}\right]$$

$$= 1 - 0.8795$$

$$= 0.1205$$

DISCRETE UNIFORM DISTRIBUTION

- A random variable X has a **discrete uniform distribution** if each of the n values in it range, say, $x_1,\ x_2, \dots, x_n$, has equal probability.
- Then,

$$P(X = x_i) = \frac{1}{n}$$

for $i = 1, 2, 3, \dots, n$.

Example 4.11

- Consider modelling the toss of a die. The outcomes are: $\{\mathbf{1, 2, 3, 4, 5, 6}\}$

Each of these outcomes are equally likely to occur and hence **discretely uniform**.

The probability for each outcome is defined as:

$$P(X = d_i) = \frac{1}{6}$$

for $d_i = 1, 2, 3, 4, 5, 6$.

Expectation, Second Moment and Variance of the Discrete Uniform

-
- Mean = $E(X) = \frac{1}{n}(total\ sum\ of\ outcomes\ or\ values)$
- Second moment =

$$E\left(X^2\right) = \frac{1}{n}(total\ sum\ of\ squared\ outcomes\ or\ squared\ values)$$

- Variance = $\sigma^2 = Var(X) = E\left(X^2\right) - [E(X)]^2$

Using the die example 4.11 above:

- Mean $= \frac{1}{6}(1+2+3+4+5+6) = \frac{15}{6} = 2.5$
- Second moment $= \frac{1}{6}\left(1^2+2^2+3^2+4^2+5^2+6^2\right) = \frac{55}{6}$
- Variance $= \text{Second moment} - \text{Mean}^2$
- Variance $= \frac{55}{6} - 2.5^2 = \frac{35}{12}$

THE GEOMETRIC DISTRIBUTION

- Suppose you flip a coin several times. **What is the probability that the first head appears on the third toss?** In order to answer this question and other similar probability questions, the geometric distribution can be used.

Conditions for a Geometric Model

For a situation to be described using a **geometric model**,

- Independent trials are carried out.
- The outcome of each trial is deemed either a success or a failure.
- The probability, p, of a successful outcome is the same for each trial.

DISTRIBUTION OF GEOMETRIC PROBABILITY

- Assume the random variable, X, **is the number of trials until a first success**, then X is said to follow a **Geometric Distribution**.
- If $X \sim Geo(p)$, the probability that the first success is obtained at the xth attempt is:

$$P(X = x) = pq^{x-1}$$

for $x = 1, 2, 3, 4, \ldots$

EXPECTATION AND VARIANCE OF THE GEOMETRIC

If $X \sim Geo(p)$,

- $E(X) = \frac{1}{p}$
- $Var(X) = \frac{q}{p^2}$

Example 4.12

- The random variable X is $Geo(p = 0.35)$.
  a) Find $P(X = 4)$.
  b) Find $P(X > 4)$.
  c) Find $P(X \leq 3)$.
  d) Find the $E(X)$.

Solution 4.12

$X \sim Geo(0.35)$.

$p = 0.35$ and $q = 1 - 0.35 = 0.65$

$P(X = x) = pq^{x-1}$.

a) $P(X = 4) = 0.35 \times 0.65^3 = 0.0961$

b) $P(X > 4) = 1 - P(X \leq 4)$

$$= 1 - [0.35 \times 0.65^0 + 0.35 \times 0.65^1 + 0.35 \times 0.65^2 + 0.35 \times 0.65^3]$$
$$= 1 - 0.8215$$
$$= 0.1785$$

c) $P(X \leq 3) = 0.35 \times 0.65^0 + 0.35 \times 0.65^1 + 0.35 \times 0.65^2$
$$= 0.7254$$

d) $E(X) = \frac{1}{p} = \frac{1}{0.35} = 2.8571$

End of Slides

Thank You