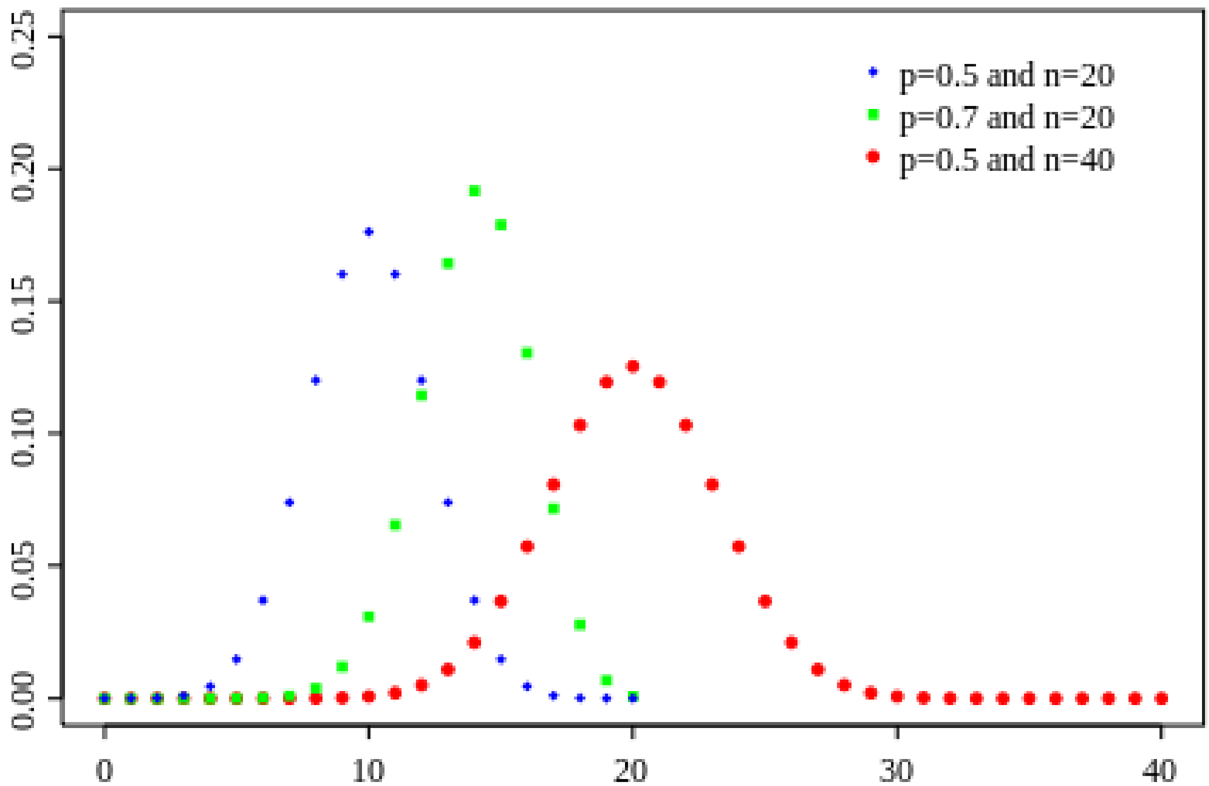
Binomial and Poisson Distribution

Dr. Gautam Kumar Meghwanshi
Dept of Microbiology
MGS University

Introduction

- A **binomial distribution** can be thought of as simply the probability of a SUCCESS or FAILURE outcome in an experiment or survey that is repeated multiple times.
- The binomial is a type of distribution that has **two possible outcomes** (the prefix “bi” means two, or twice).
- For example, a coin toss has only two possible outcomes: heads or tails and taking a test could have two possible outcomes: pass or fail.

- The first variable in the binomial formula, **n**, stands for the number of times the experiment runs.
- The second variable, **p**, represents the probability of one specific outcome.
- For example, let's suppose you wanted to know the probability of getting a 1 on a die roll.
- If you were to roll a die 20 times, the probability of rolling a one on any throw is 1/6.
- Roll twenty times and you have a binomial distribution of (n=20, p=1/6). SUCCESS would be "roll a one" and FAILURE would be "roll anything else." If the outcome in question was the probability of the die landing on an even number, the binomial distribution would then become (n=20, p=1/2). That's because your probability of throwing an even number is one half.



A Binomial Distribution shows either (S)uccess or (F)ailure.

Criteria

- Binomial distributions must also meet the following three criteria:
- **The number of observations or trials is fixed.** In other words, you can only figure out the probability of something happening if you do it a certain number of times. This is common sense—if you toss a coin once, your probability of getting a tails is 50%. If you toss a coin a 20 times, your probability of getting a tails is very, very close to 100%.
- **Each observation or trial is** independent. In other words, none of your trials have an effect on the probability of the next trial.
- The **probability of success** (tails, heads, fail or pass) is **exactly the same** from one trial to another.

- Once you know that your distribution is binomial, you can apply the **binomial distribution formula** to calculate the probability.

What is a Binomial Distribution?

- Many instances of binomial distributions can be found in real life.
- For example, if a new drug is introduced to cure a disease, it either cures the disease (it’s successful) or it doesn’t cure the disease (it’s a failure).
- If you purchase a lottery ticket, you’re either going to win money, or you aren’t.
- Basically, anything you can think of that can only be a success or a failure can be represented by a binomial distribution.

The Binomial Distribution Formula

- The binomial distribution formula is:
- **$b(x; n, P) = {}_nC_x * P^x * (1 - P)^{n - x}$**
- Where:
  b = binomial probability
  x = total number of "successes" (pass or fail, heads or tails etc.)
  P = probability of a success on an individual trial
  n = number of trials

- **Note:** The binomial distribution formula can also be written in a slightly different way, because **$_nC_x$ = n! / x!(n – x)!** (this binomial distribution formula uses factorials.
- $$P(X) = \frac{n!}{(n-X)!\,X!} \cdot (p)^X \cdot (q)^{n-X}$$
- “q” in this formula is just the probability of failure (subtract your probability of success from 1).

Using the First Binomial Distribution Formula

- The binomial distribution formula can calculate the probability of success for binomial distributions.
- **Example 1**
- **Q. A coin is tossed 10 times. What is the probability of getting exactly 6 heads?**
- We can use this formula:
- $b(x; n, P) = {}_{n}C_{x} * P^{x} * (1 – P)^{n – x}$
  The number of trials (n) is 10
  The odds of success (“tossing a heads”) is 0.5 (So 1-p = 0.5)
  x = 6
- $P(x=6) = {}_{10}C_{6}$ * 0.5^6 * 0.5^4 = 210 * 0.015625 * 0.0625 = 0.205078125

- **Example 2**

$$P(X) = \frac{n!}{(n-X)!\,X!} \cdot (p)^X \cdot (q)^{n-X}$$

- **80% of people who purchase pet insurance are women.**
  If 9 pet insurance owners are randomly selected, find the probability that exactly 6 are women.
- **Step 1:** Identify 'n' from the problem. Using our example question, n (the number of randomly selected items) is 9.
- **Step 2:** Identify 'X' from the problem. X (the number you are asked to find the probability for) is 6.
- **Step 3:** Work the first part of the formula. The first part of the formula is
- n! / (n – X)! X!
- Substitute your variables:
- 9! / ((9 – 6)! × 6!)
- Which equals **84**. Set this number aside for a moment.

- **Step 4:** Find p and q. p is the probability of success and q is the probability of failure. We are given p = 80%, or 0.8. So the probability of failure is 1 – 0.8 = 0.2 (20%).
- **Step 5:** Work the second part of the formula.
- p^X
  $= 0.8^6$
  $= 0.262144$
- Set this number aside for a moment.
- **Step 6:** Work the third part of the formula.
- $q^{(n - X)}$
  $= 0.2^{(9-6)}$
  $= 0.2^3$
  $= 0.008$
- **Step 7:** Multiply your answer from step 3, 5, and 6 together.
  $84 \times 0.262144 \times 0.008 = 0.176.$

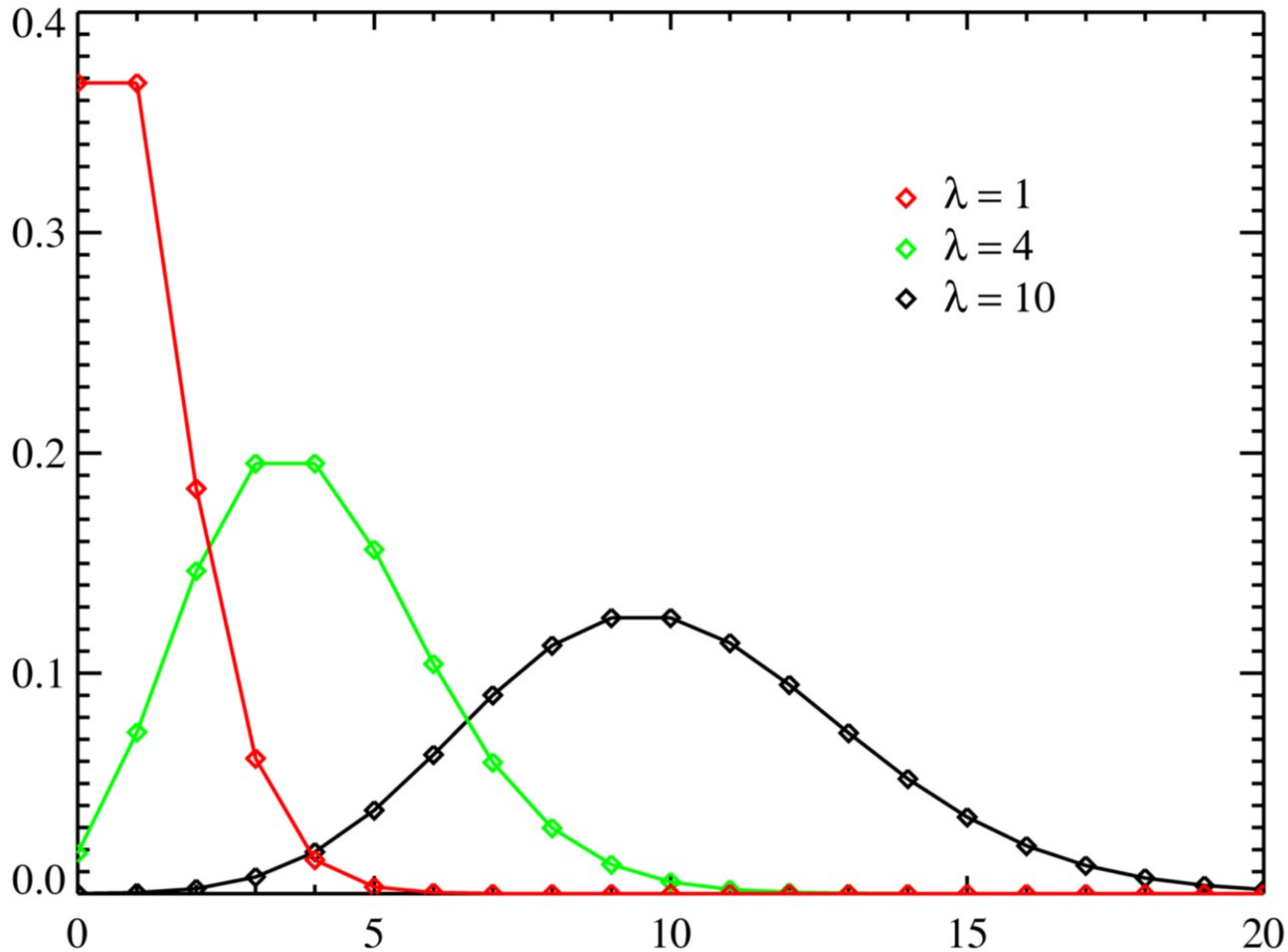
- **Example 3**
- **60% of people who purchase sports cars are men. If 10 sports car owners are randomly selected, find the probability that exactly 7 are men.**
- **Step 1:**: Identify 'n' and 'X' from the problem. Using our sample question, n (the number of randomly selected items—in this case, sports car owners are randomly selected) is 10, and X (the number you are asked to "find the probability" for) is 7.
- **Step 2:** Figure out the first part of the formula, which is:
- n! / (n – X)! X!
- Substituting the variables:
- 10! / ((10 – 7)! × 7!)
- Which equals 120. Set this number aside for a moment.

- **Step 3:** Find “p” the probability of success and “q” the probability of failure. We are given p = 60%, or 0.6. therefore, the probability of failure is 1 – 0.6 = 0.4 (40%).
- **Step 4:** Work the next part of the formula.
- p^X
  $= 0.6^7$
  $= 0.0279936$

  Set this number aside while you work the third part of the formula.
- **Step 5:** Work the third part of the formula.
- $q^{(n-7)}$
  $= 0.4^{(10-7)}$
  $= 0.4^3$
  $= 0.064$
- **Step 6:** Multiply the three answers from steps 2, 4 and 5 together.
  $120 \times 0.0279936 \times 0.064 = 0.215$.
- That’s it!

Poisson Distribution

- A Poisson distribution is a tool that helps to predict the probability of certain events from happening when you know how often the event has occurred.
- It gives us the **probability of a given number of events happening in a fixed interval of time**.



Poisson distributions, valid only for integers on the horizontal axis. λ (also written as μ) is the expected number of event occurrences.

Practical Uses of the Poisson Distribution

- A textbook store rents an average of 200 books every Saturday night. Using this data, you can **predict the probability that more books will sell** (perhaps 300 or 400) on the following Saturday nights.
- Another example is the number of diners in a certain restaurant every day. If the average number of diners for seven days is 500, you can predict the probability of a certain day having more customers.

- Because of this application, Poisson distributions are used by businessmen to make **forecasts** about the number of customers or sales on certain days or seasons of the year.
- In business, overstocking will sometimes mean losses if the goods are not sold. Likewise, having too few stocks would still mean a lost business opportunity because you were not able to maximize your sales due to a shortage of stock.
- By using this tool, businessmen are able to estimate the time when demand is unusually higher, so they can purchase more stock.
- Hotels and restaurants could prepare for an influx of customers, they could hire extra temporary workers in advance, purchase more supplies, or make contingency plans just in case they cannot accommodate their guests coming to the area.

- With the Poisson distribution, companies can adjust supply to demand in order to keep their business earning good profit. In addition, waste of resources is prevented.

Calculating the Poisson Distribution

- *The Poisson Distribution pmf is:*

  $P(x; \mu) = (e^{-\mu} * \mu^{x}) / x!$
- Where:
- The symbol "!" is a factorial.
- μ (the expected number of occurrences) is sometimes written as λ. Sometimes called the **event rate** or rate parameter.

- **Example question**
- **The average number of major storms in your city is 2 per year. What is the probability that exactly 3 storms will hit your city next year?**

Step 1: Figure out the components you need to put into the equation.

- $\mu = 2$ (average number of storms per year, historically)
- $x = 3$ (the number of storms we think might hit next year)
- $e = 2.71828$ (e is Euler's number, a constant)

Step 2: Plug the values from Step 1 into the Poisson distribution formula:

- $P(x; \mu) = (e^{-\mu}) (\mu^{x}) / x!$

  $= (2.71828^{-2}) (2^{3}) / 3!$

  $= (0.13534) (8) / 6$

  $= 0.180$

- The probability of 3 storms happening next year is 0.180, or 18%

Poisson distribution vs. Binomial

- If your question has an **average probability** of an event happening per unit (i.e. per unit of time, cycle, event) **and** you want to find probability of a certain number of events happening in a period of time (or number of events), then use the Poisson Distribution.
- If you are given an **exact probability** and you want to find the probability of the event happening a certain number out times out of x (i.e. 10 times out of 100, or 99 times out of 1000), use the **Binomial Distribution formula**.