Discrete Random Variable

Binomial Distribution:

- All the trails are independent
- Number (n) of trails is finite
- The Probability (p) of the success is same of each trials

$$P(x) = C_x^n \, p^x q^{(n-x)}$$

Discrete Random Variable

Binomial Distribution:

- All the trails are independent
- Number (n) of trails is finite
- The Probability (p) of the success is same of each trials

$$P(x) = C_x^n \, p^x q^{(n-x)}$$

Example:

a. A coin toss 3-times, find the probability of 2-Heads.

b. A coin toss 10-times, find the probability of 5-Heads.

Binomial Distribution

Q1. The probability that man aged 60 will live up to 70 is 0.65 out of 10 men. Now aged 60, find the probability:

1. At least 7 will live up to 70

2. Exactly 9 will live up to 70

3. At most 9 will live up to 70

Binomial Distribution

Q2. Out of 800 families with 5 children each, how many families would be expected to have

- 3 boys
- 5 girls
- Either 2 or 3 boys
- At least 2 girls

Binomial Distribution

Q3. The Probability that a pen manufactured by a company will be defective is 1/10. If 12 such pen are manufactured. Find the probability that:

1. Exactly 2 will be defective
2. None will be defective
3. At least 2 will be defective

Binomial Distribution

Q4. Medical professionals use the binomial distribution to model the probability that a certain number of patients will experience side effects as a result of taking new medications.

E.g., suppose it is known that 5% of adults who take a certain medication experience negative side effects. We can use a Binomial Distribution Calculator to find the probability that more than a certain number of patients in a random sample of 100 will experience negative side effects.

- P(X > 5 patients experience side effects) = **??**
- P(X > 10 patients experience side effects) = **??**
- P(X > 15 patients experience side effects) = **??**

Binomial Distribution

Q4. Medical professionals

suppose it is known that 5% of adults who take a certain medication experience negative side effects. Find the probability that more than a certain number of patients in a random sample of 100 will experience negative side effects.

- P(X > 5 patients experience side effects) = **0.38400**
- P(X > 10 patients experience side effects) = **0.01147**
- P(X > 15 patients experience side effects) = **0.0004**

$p = 0.05$ $\qquad p(x) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$q = 0.95$

$n = 100$

$P(X > 5) = 1 - P(X \leq$

$= 1 - [P(0) + P(1) + P(2) + P(3) +$

Binomial Distribution

Q5. Banks use the binomial distribution to model the probability that a certain number of credit card transactions are fraudulent.

E. g., suppose it is known that 2% of all credit card transactions in a certain region are fraudulent. If there are 50 transactions per day in a certain region, we can use a Binomial Distribution Calculator to find the probability that more than a certain number of fraudulent transactions occur in a given day:

- P(X > 1 fraudulent transaction) = ??
- P(X > 2 fraudulent transactions) = **??**
- P(X > 3 fraudulent transactions) = **??**

Binomial Distribution

Q5. Banks use the binomial distribution to model the probability that a certain number of credit card transactions are fraudulent.

E. g., suppose it is known that 2% of all credit card transactions in a certain region are fraudulent. If there are 50 transactions per day in a certain region, we can use a Binomial Distribution Calculator to find the probability that more than a certain number of fraudulent transactions occur in a given day:

- P(X > 1 fraudulent transaction) = **0.26423**
- P(X > 2 fraudulent transactions) = **0.07843**
- P(X > 3 fraudulent transactions) = **0.01776**

$p = 0.02$ $\quad n = 50$

$q = 0.98$ $\quad p(x) = {}^{n}C_{x}\, p^{x} q^{n-x}$

$P(X > 1) = 1 - [P(0) + P(1)]$

Negative Binomial Distribution

NBD is applicable when we need to performed an experiment untill a total of r success are obtained

Note: If r = 1, means we perform an experiment till we obtained first success.

Negative Binomial Distribution

Take a standard deck of cards, shuffle them, and choose a card. Replace the card and repeat until you have drawn two Kings.

Y is the number of draws needed to draw two Kings.

As the number of trials isn't fixed (i.e. you stop when you draw the second King), this makes it a negative binomial distribution.

Negative Binomial Distribution

$$P(x) = \left(C_{r-1}^{x-1}\, p^{r-1} q^{(x-1)-(r-1)}\right) p$$

$$P(x) = {}^{n}C_{x} \cdot p^{x} \cdot q^{n-x}$$

Negative Binomial Distribution

Q1. If the probability is 0.40 that a child exposed the certain disease will contain it. What is the probability that the 10th child exposed to the disease will be the 3rd to catch ?

$$P(X=10) = \left({}^{9}C_{2} \cdot p^{2} \cdot q^{7}\right) p$$

$$= \frac{\lfloor 9}{\lfloor 7 \lfloor 2} (0.4)^{2} (.6)^{7} \times 0$$

$$= 0.064$$

Negative Binomial Distribution

Q3. Let x be the number of births in a family until the 2nd daughter is born. If the probability of the having a male child is ½ . Find the probability that the 6th child in the family is the second daughter.

Bernoulli Distribution

A discrete random variable X is said to have a Bernoulli distribution with parameter p. If its probability mass function is given by:

$$P(x) = p^{x}(1-p)^{n-x} \quad , x = 0, 1$$

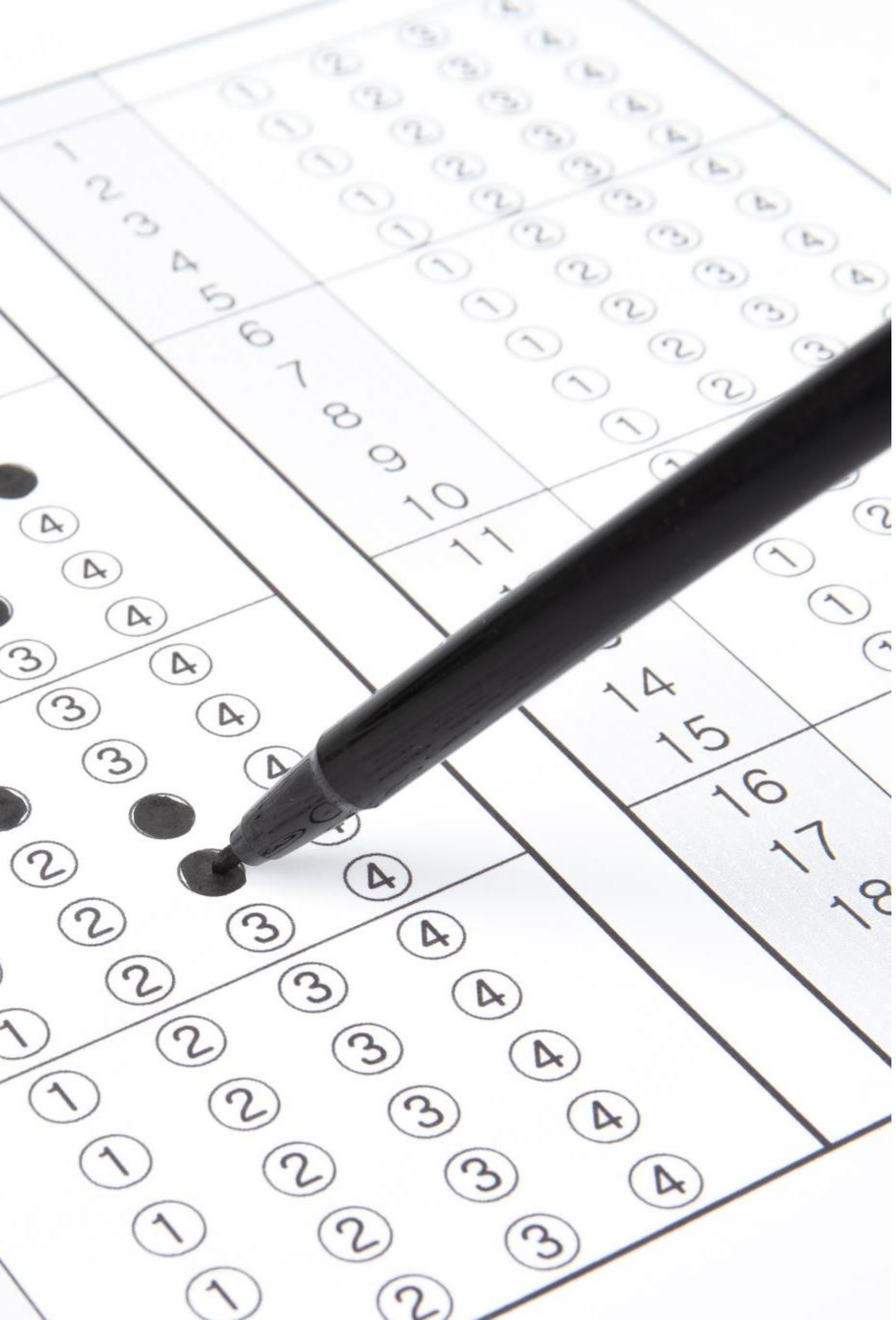
Bernoulli distribution

Bernoulli distribution arises when the following 3-conditions are satisfied.

1. Each trail of an experiment results in an outcome that may be classified as a success or failure
2. The probability of a success $P(S) \ = \ p$ is the same for each trail.
3. The trails are independent; that is the outcome of one trail have no effect on the outcome of any other trail.

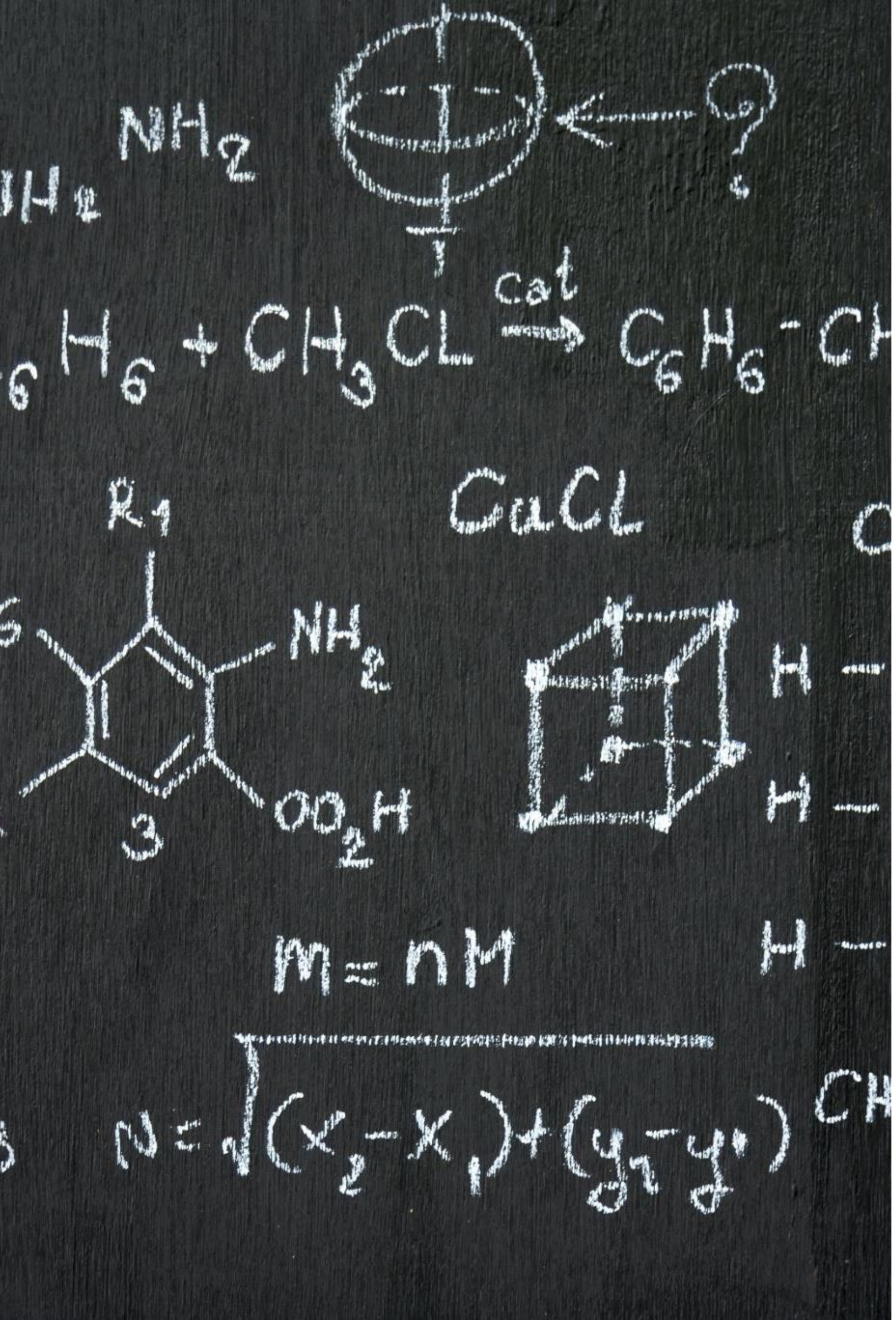
Bernoulli distribution

- The prevalence of a certain disease in the general population is 10%.
- If we randomly select a person from this population, we can have only two possible outcomes (diseased or healthy person). We call one of these outcomes (diseased person) success and the other (healthy person), a failure.
- The probability of success (p) or diseased person is 10% or 0.1. So, the probability of failure (q) or healthy person = 1-p = 1-0.1 = 0.9.



Bernoulli distribution

In an exam, 10 multiple choice questions are asked where only one out of four questions are correct. Find the probability of getting 5 out of 10 questions correct in an answer sheet.

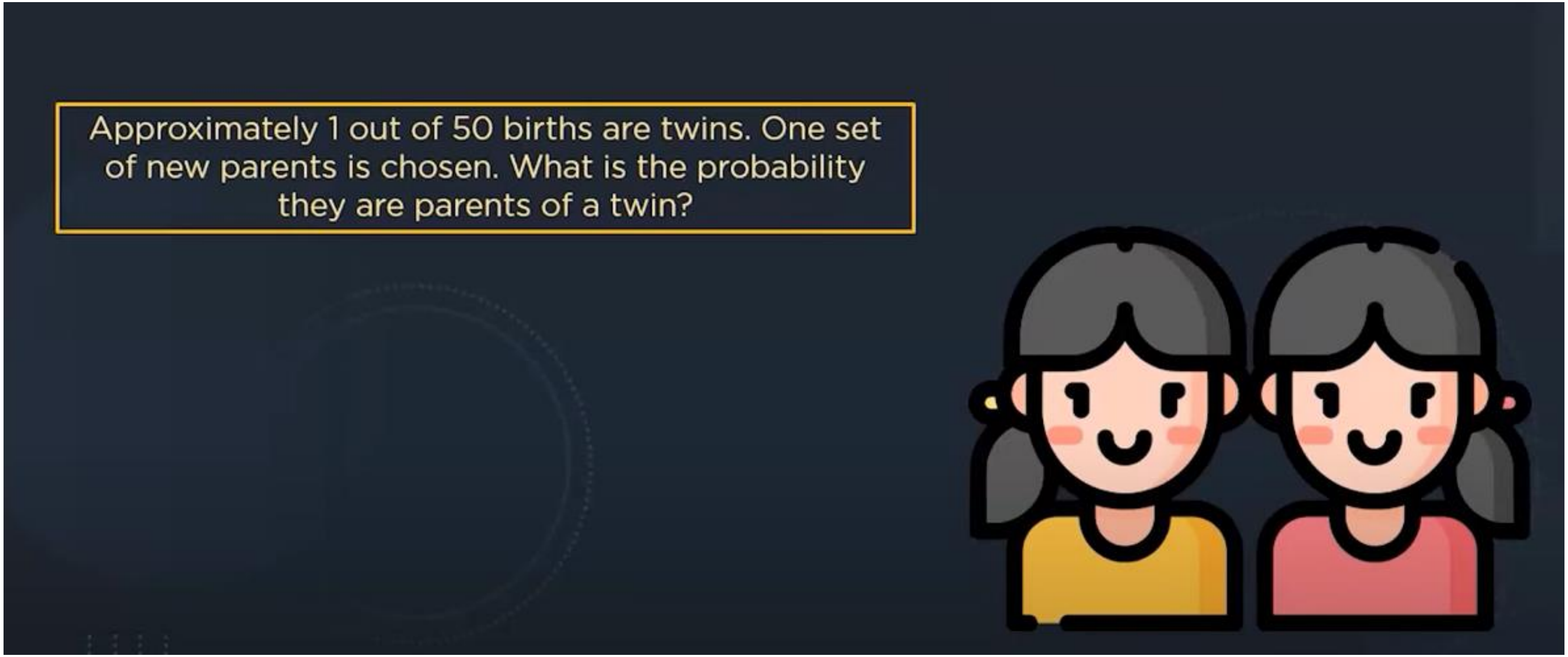


Bernoulli distribution

Solution:

Probability of getting an answer correct, p = ¼
Probability of getting an answer incorrect , q = 1-p = 1
Probability of getting 5 answers correct, P(X=5) = (0.25)5 (0.75)5 = 0.5839920044

Bernoulli distribution



Bernoulli distribution

Approximately 1 out of 50 births are twins. One set of new parents is chosen. What is the probability they are parents of a twin?

$$P = 1/50$$

$$P(X=1) = (1/50)^1 * (1-1/50)^0$$
$$= 1/50$$

Bernoulli distribution

Real-World Examples

- Whether a person voted for a particular political party or abstained from voting. Success = 1 if voted for a political party

Bernoulli distribution

- Whether the outcome of an interview is recommendation for next round of interview or not. Success = 1 if the result of interview is recommendation to the next round of interview

Exercise Problems

Q1. You are surveying people exiting from a polling booth and asking them if they voted independent. The probability (p) that a person voted independent is 20%. What is the probability that 15 people must be asked before you can find 5 people who voted independent?

Exercise Problems

Q3. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil.

What is the probability that the first strike comes on the third well drilled?

Exercise Problems

Q4. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil.

What is the probability that the third strike comes on the seventh well drilled?

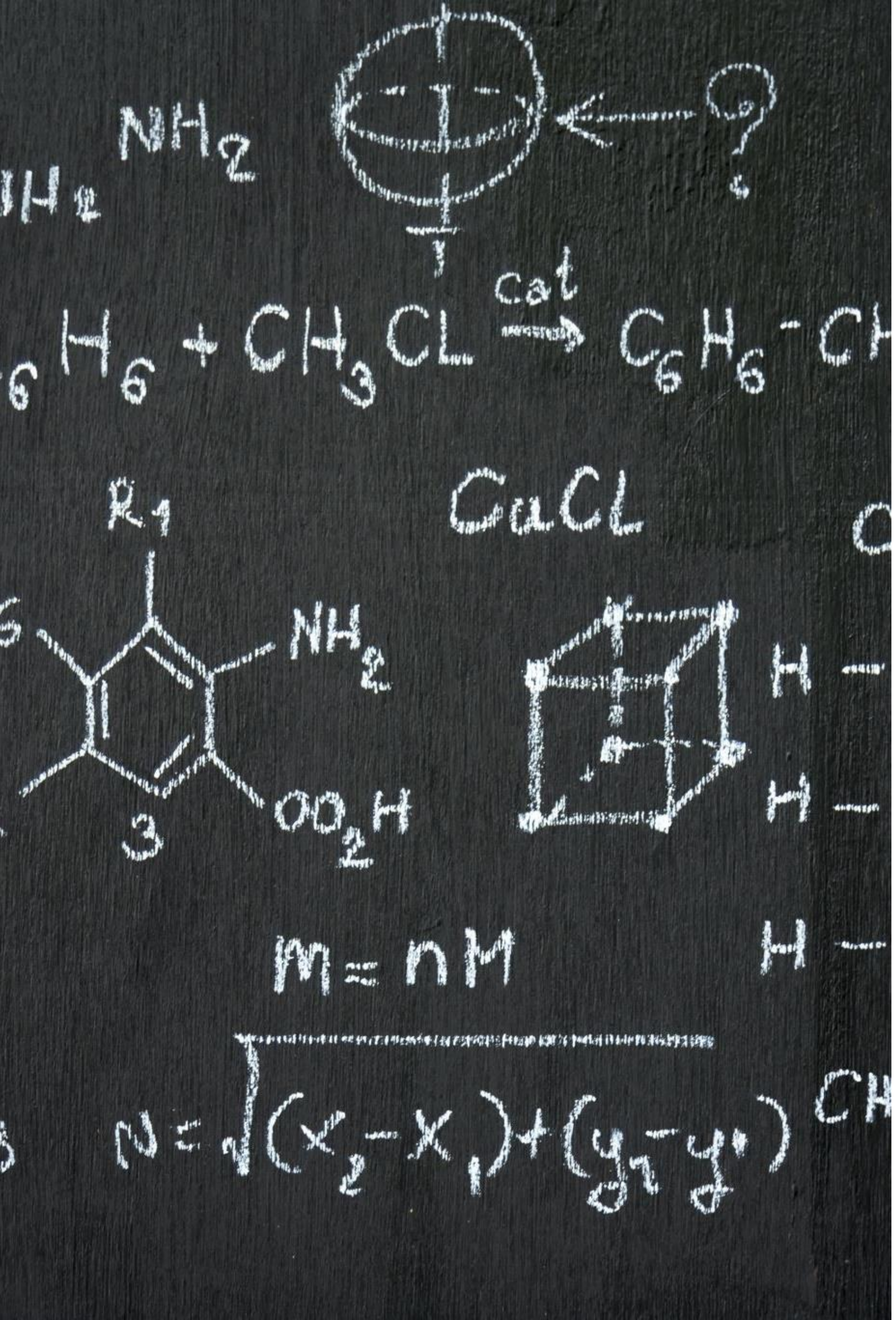
Exercise Problems

Q5. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil.

What is the mean and variance of the number of wells that must be drilled if the oil company wants to set up three producing wells?

Exercise Problems

Q6. Bob is a high school basketball player. He is a 70% free throw shooter. That means his probability of making a free throw is 0.70. During the season, what is the probability that Bob makes his third free throw on his fifth shot?



Q6. Solution

This is an example of a negative binomial experiment. The probability of success (P) is 0.70, the number of trials (x) is 5, and the number of successes (r) is 3.

We enter these values into the negative binomial formula.

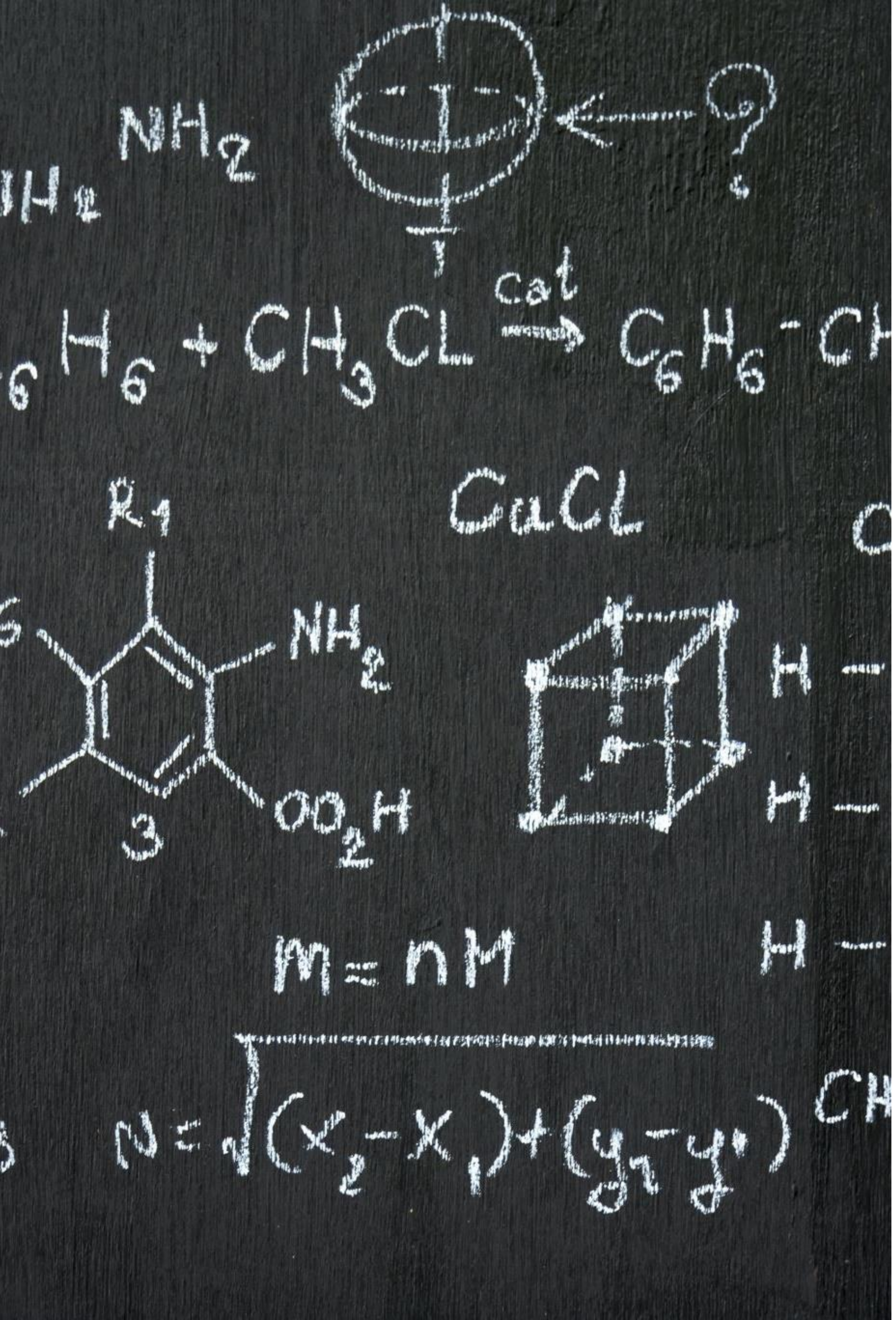
$b^*(x; r, P) = {}_{x-1}C_{r-1} * P^r * Q^{x - r}$
$b^*(5; 3, 0.7) = {}_4C_2 * 0.7^3 * 0.3^2$
$b^*(5; 3, 0.7) = 6 * 0.343 * 0.09 = 0.18522$

Thus, the probability that Bob will make his third successful free throw on his fifth shot is 0.18522.

Exercise Problems

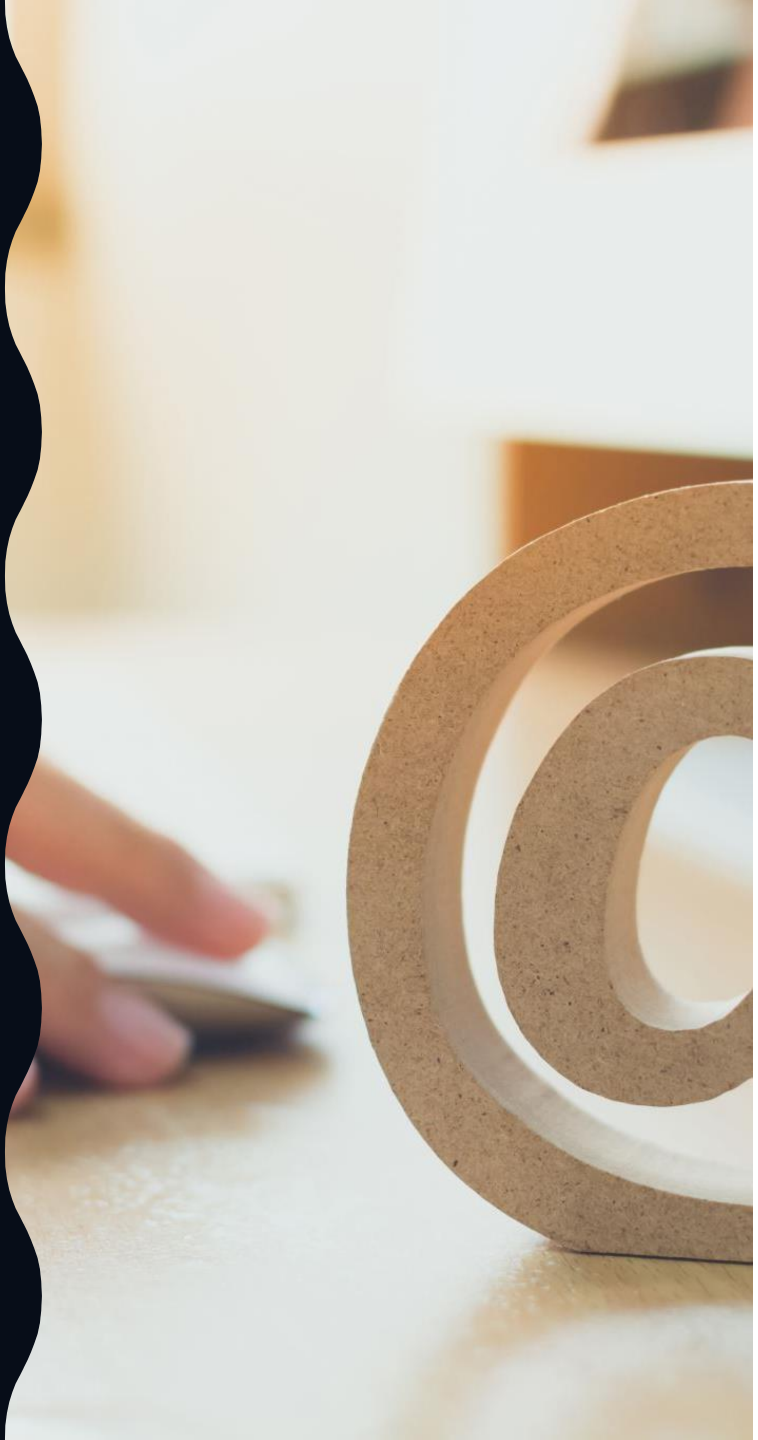
Q7. Bob is a high school basketball player. He is a 70% free throw shooter. That means his probability of making a free throw is 0.70. During the season,

What is the probability that Bob makes his first free throw on his fifth shot?



Q7. Solution

- The probability of success (*P*) is 0.70, the number of trials (*x*) is 5, and the number of successes (*r*) is 1. We enter these values into the negative binomial formula.
  - $b^*(x; r, P) = {}_{x-1}C_{r-1} * P^r * Q^{x - r}$
    $b^*(5; 1, 0.7) = {}_4C_0 * 0.7^1 * 0.3^4$
    $b^*(5; 3, 0.7) = 0.00567$



Exercise Problems

Q8. **Number of Spam Emails per Day**

Suppose it is known that 4% of all emails are spam. If an account receives 20 emails in a given day, find the probability that a certain number of those emails are spam:

- P(X = 0 spam emails)
- P(X = 1 spam email)
- P(X = 2 spam emails)

Exercise Problems

Q8. Suppose it is known that 4% of all emails are spam. If an account receives 20 emails in a given day, find the probability that a certain number of those emails are spam:

- P(X = 0 spam emails) = **0.44200**
- P(X = 1 spam email) = **0.36834**
- P(X = 2 spam emails) = **0.14580**

Exercise Problems

Q9. river overflows

suppose it is known that a given river overflows during 5% of all storms. If there are 20 storms in a given year, find the probability that the river overflows a certain number of times:

P(X = 0 overflows)

P(X = 1 overflow)

P(X = 2 overflows)

Exercise Problems

Q9. River overflows

suppose it is known that a given river overflows during 5% of all storms. If there are 20 storms in a given year, find the probability that the river overflows a certain number of times:

- P(X = 0 overflows) = **0.35849**
- P(X = 1 overflow) = **0.37735**
- P(X = 2 overflows) = **0.18868**

Poisson Distribution

- Suppose we are counting the number of occurrences of an event in a given unit of time, distance, area, or volume.

For Example:

- The number of car accidents in a day
- The number of defective items in a lot

Poisson Distribution

Suppose:

- Events are occurring independently
- The probability that an event occurs in a given length of time does not change through time.

- Then X, the number of events in a fixed unit of time, has a Poisson Distribution

Poisson Distribution

- Then X, the number of events in a fixed unit of time, has a Poisson Distribution

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}; x = 0, 1, \dots.$$

Poisson Distribution

$p = 0.02$

$n = 200$

Q1. Given that 2% at the fuses manufactured by a firm are defective. Find probability that a box containing 200 fuses has

1. At least 1 defective fuses
2. 3 or more defective fuses
3. No defective fuses

$\lambda =$

$=$

$\lambda =$

$P(0) = \dfrac{\lambda^0 \cdot e^{-4}}{0!}$

$P(x \geq 1) \Rightarrow 1 - P($

$\Rightarrow 1 - P($

Poisson Distribution

Solution:

Sol'n $n = 200$, $p = 0.02$

$$\boxed{\lambda = np = 200(0.02) = 4}$$

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(4)^x e^{-4}}{x!}$$

(i) $P(X \geq 1) = 1 - P(X < 1)$

$$= 1 - P(0)$$

$$= 1 - \frac{4^0 e^{-4}}{0!} = 1 - e^{-4}$$

(ii) $P(X \geq 3) = 1 - P(X < 3)$

$$= 1 - [P(0) + P(1) + P(2)]$$

$$= 1 - \left(\frac{4^0 e^{-4}}{0!} + \frac{4^1 e^{-4}}{1!} + \frac{4^2 e^{-4}}{2!}\right)$$

$$= 1 - e^{-4}(1 + 4 + 8) = 1 - 13e^{-4}$$

(iii) $P(0) = \frac{(4)^0 e^{-4}}{0!} = e^{-4}$

Poisson Distribution

Q2. A certain factory turning out blades there is small chance of 0.002 for any blade to be defective.

The blades are supplied in packets of 10.

Find approximate number of packets containing

1. No defective blades

2. One defective blades

In Consignment of 10000 packets

Poisson Distribution

Q2. A certain factory turning out blades there is small chance of 0.002 for any blade to be defective.

The blades are supplied in packets of 10. Using poisson distributions. Find approximate number of packets containing

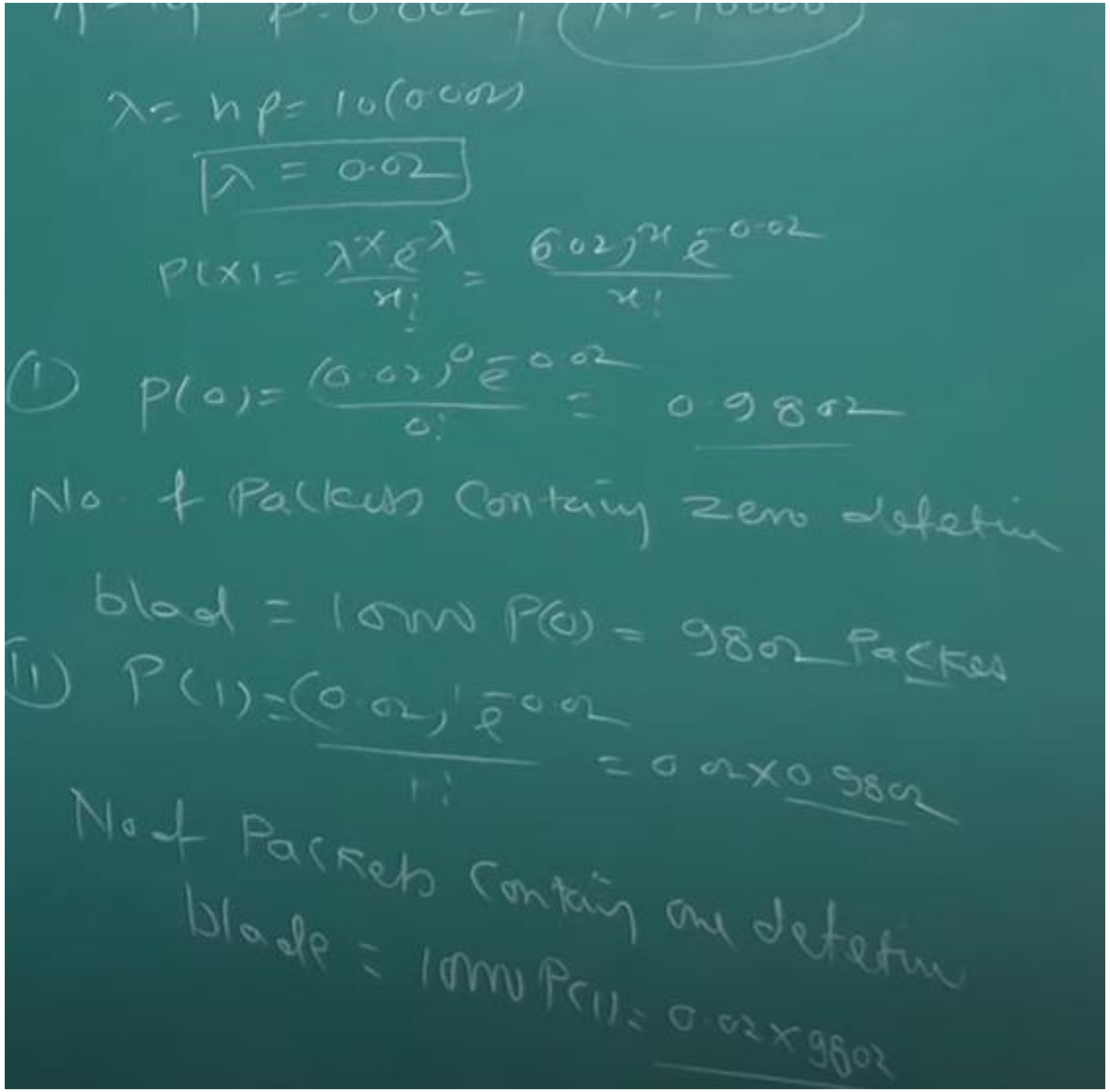
1. No defective blades

2. One defective blades

In Consignment of 10000 packets

Solution:

Poisson Distribution



Poisson Distribution

Q3. If probability of a bad reaction from a certain infection is 0.01.

Find the chance that out of 200 individuals more than two will get bad reaction.

Poisson Distribution

Solution:

Sol^n

$p = 0.01, \quad n = 200$

$\lambda = np = (0.01)(200) = 2$

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(2)^x e^{-2}}{x!}$$

$$P(x > 2) = 1 - P(x \le 2)$$

$$= 1 - (P(0) + P(1) + P(2))$$

$$= 1 - \left(\frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!} \right)$$

$$= 1 - e^{-2}(1 + 2 + 2)$$

$$= 1 - 5e^{-2}$$

The Relationship Between the Binomial and Poisson Distributions

The binomial distribution tends toward the Poisson distribution as: $n \to \infty$, $p \to 0$, and $\lambda = np$ stays constant.

Binomial:

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

$\lambda = np$

$p = \lambda / n$

Poisson:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

The Poisson distribution can be used to provide a reasonable approximation to the binomial distribution if n is large and p is small.

Example:

Albinism is a rare genetic disorder that affects one in 20,000 Europeans.

People with albinism produce little or none of the pigment melanin.

Example:

Albinism is a rare genetic disorder that affects one in 20,000 Europeans.

People with albinism produce little or none of the pigment melanin.

In a random sample of 1000 Europeans, what is the probability that exactly 2 have albinism?

Solution

Binomial $n=1000$ $p = \frac{1}{20000}$

$$P(x=2) = \binom{1000}{2}\left(\frac{1}{20000}\right)^2\left(1-\frac{1}{20000}\right)^{1000-2}$$

$$= \underline{0.001187965}$$

Poisson $\frac{\lambda^x e^{-\lambda}}{x!}$

$$\lambda = np = 1000 \cdot \frac{1}{20000} = 0.05$$

$$P(x=2) = \frac{0.05^2 e^{-0.05}}{2!}$$

$$= 0.001187$$

Rough Guideline:

The Poisson approximation is reasonable if $n > 50$ and $np < 5$.

Why use this approximation?

- The factorials and exponentials in the binomial formula can become problematic to calculate.
- A problem may be binomial conceptually, but n and p may be unknown. (We may only know the mean.)

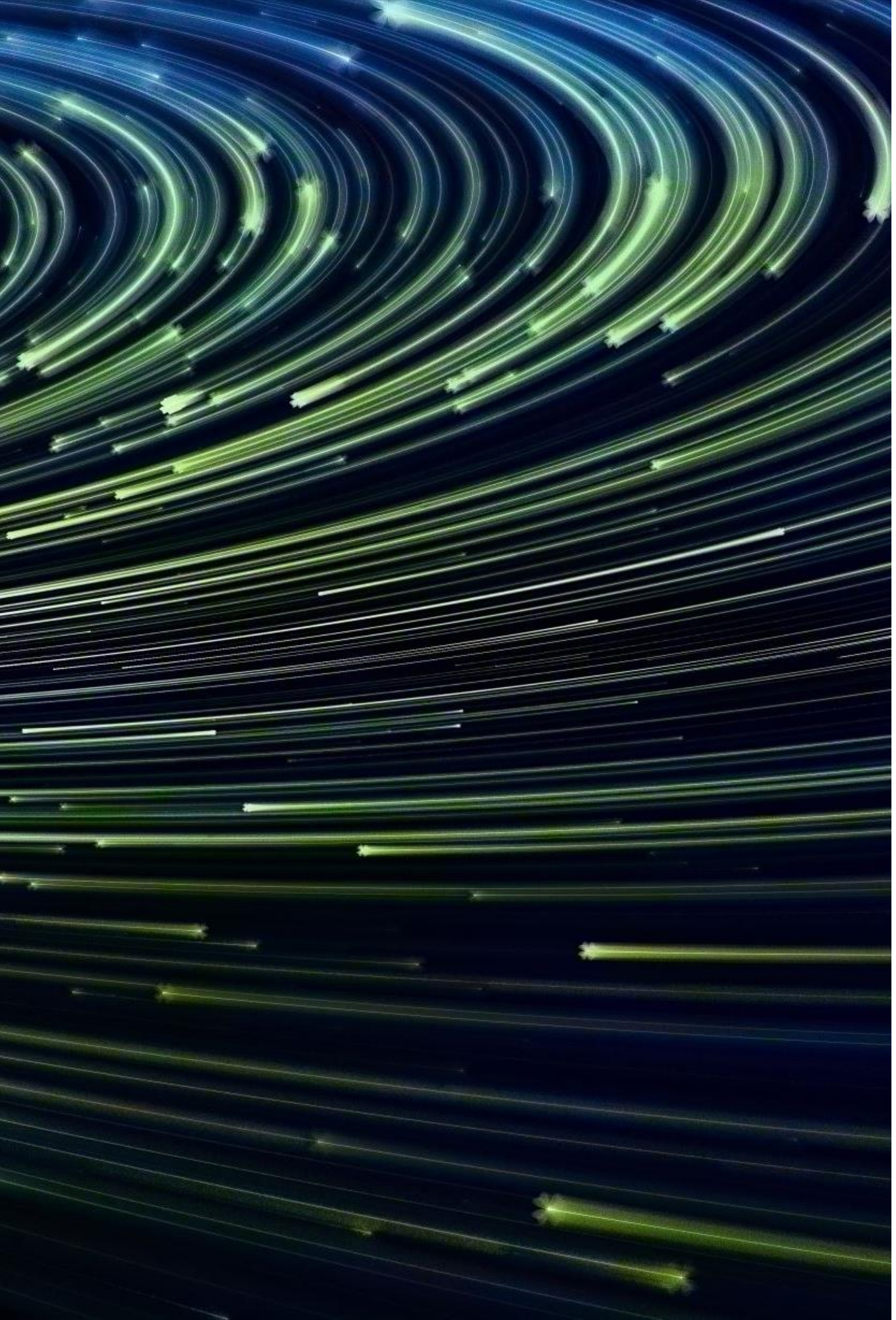
Poisson Distribution

Problem: Network Failures

- A tech company wants to model the probability that a certain number of network failures occur in a given week.
- Suppose it's known that an average of 4 network failures occur each week.
- Let *X* be the number of network failures in a given week.
- What type of distribution does the random variable *X* follow?

Poisson Distribution

- **Problem: Network Failures**
- A tech company wants to model the probability that a certain number of network failures occur in a given week. Suppose it's known that an average of 4 network failures occur each week. Let *X* be the number of network failures in a given week. What type of distribution does the random variable *X* follow?
- Answer: *X* follows a Poisson distribution because we're interested in modeling the number of network failures in a given week and there is no upper limit on the number of failures that could occur. This is not a Binomial distribution because there is not a fixed number of trials.



Poisson Distribution

Problem: Shooting Free-Throws

Tyler makes 70% of all free-throws he attempts.

Suppose he shoots 10 free-throws.

Let *X* be the number of times Tyler makes a basket during the 10 attempts.

What type of distribution does the random variable *X* follow?

Poisson Distribution

Problem: Shooting Free-Throws

- Tyler makes 70% of all free-throws he attempts. Suppose he shoots 10 free-throws. Let X be the number of times Tyler makes a basket during the 10 attempts. What type of distribution does the random variable X follow?
- Answer: X follows a Binomial distribution because there is a fixed number of trials (10 attempts), the probability of “success” on each trial is the same, and each trial is independent.