

Probability and Statistics

Prepared by Dr. Alfaisal Hasan

STA 201
Probability and Statistics

Lecture NO 4



Probability and Statistics

Lecture NO 4

STA 201

Probability and Statistics

Prepared by
Dr. Alfaisal Hasan

Chapter

Topics

- Chapter 1: Introduction To Statistics
- Chapter 2: Descriptive Statistics
- **Chapter 3: Probability**
- Chapter 4: Discrete Probability Distributions
- Chapter 5: Normal Probability Distributions
- Chapter 6: Confidence Intervals
- Chapter 7: Hypothesis Testing with One Sample
- Chapter 8: Hypothesis Testing with Two Samples
- Chapter 9: Correlation And Regression
- Chapter 10: Chi-Square Tests and The F -Distribution

الفصل الأول: مقدمة في الإحصاء

الفصل الثاني: الإحصاء الوصفي

الفصل الثالث: الاحتمالات

الفصل الرابع: التوزيعات الاحتمالية المنفصلة

الفصل الخامس: التوزيع الاحتمالي الطبيعي

الفصل السادس: فترات الثقة

الفصل السابع: اختبار الفرضيات لعينة واحدة

الفصل الثامن: اختبار الفرضيات لعينتين

الفصل التاسع: الارتباط والانحدار

الفصل العاشر: اختبارات كاي تربيع وتوزيع F

Discrete Probability Distributions

التوزيعات الاحتمالية المنفصلة

4.1 Probability Distributions

التوزيعات الاحتمالية

4.2 Binomial Distributions

توزيع ذي الحدين

4.3 More Discrete Probability Distributions

المزيد من التوزيعات الاحتمالية المنفصلة

Discrete Probability Distributions

- Distinguish between discrete random variables and continuous random variables
- Construct a discrete probability distribution and its graph
- Determine if a distribution is a probability distribution
- Find the mean, variance, and standard deviation of a discrete probability distribution
- Find the expected value of a discrete probability distribution

التمييز بين المتغيرات العشوائية المنفصلة والمتغيرات العشوائية المتصلة

إنشاء توزيع احتمالي منفصل ورسمه البياني

تحديد ما إذا كان التوزيع توزيعًا احتماليًا

حساب المتوسط والتباين والانحراف المعياري للتوزيع الاحتمالي المنفصل

حساب القيمة المتوقعة للتوزيع الاحتمالي المنفصل

Probability Distributions

Random Variable

Represents a numerical value associated with each outcome of a probability distribution.

Denoted by x

Examples

x = Number of sales calls a salesperson makes in one day.

x = Hours spent on sales calls in one day.

متغير عشوائي

يمثل قيمة عددية مرتبطة بكل نتيجة للتوزيع الاحتمالي.

يُشار إليه بـ X
أمثلة

X = عدد مكالمات المبيعات التي يجريها مندوب المبيعات في يوم واحد.

X = الساعات المستغرقة في مكالمات المبيعات في يوم واحد.

Discrete Random Variable

المتغير العشوائي المنفصل

Discrete Random Variable

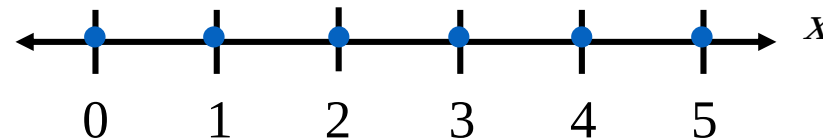
Has a finite or countable number of possible outcomes that can be listed.

لديه عدد محدود أو لا يحصى من النتائج المحتملة التي يمكن سردها.

Example:

x = Number of sales calls a salesperson makes in one day.

عدد مكالمات المبيعات التي يجريها مندوب المبيعات في يوم واحد.



Continuous Random Variable

المتغير العشوائي المتصل

Continuous Random Variable

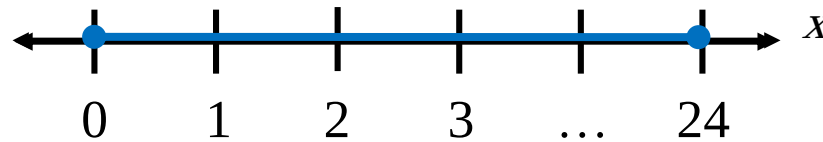
Has an uncountable number of possible outcomes, represented by an interval on the number line.

لديه عدد لا يحصى من النتائج المحتملة، ممثلة بفترة على خط الأعداد.

Example:

x = Hours spent on sales calls in one day.

الساعات التي يتم قضاؤها في مكالمات المبيعات في يوم واحد.



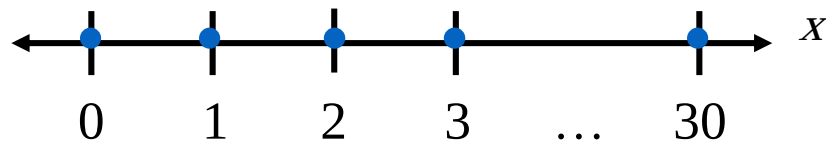
Example: Random variable

Decide whether the random variable x is discrete or continuous.

x = The number of stocks in the Dow Jones Industrial Average that have share price increases on a given day.

Solution:

Discrete random variable (The number of stocks whose share price increases can be counted.)



حدد ما إذا كان المتغير العشوائي منفصلاً أم متصلاً؟
عدد الأسهم في مؤشر داو جونز الصناعي التي ترتفع أسعار أسهمها في
يوم معين.

الحل:

المتغير العشوائي المنفصل (عدد الأسهم التي يمكن حساب الزيادات في
أسعار أسهمها.)

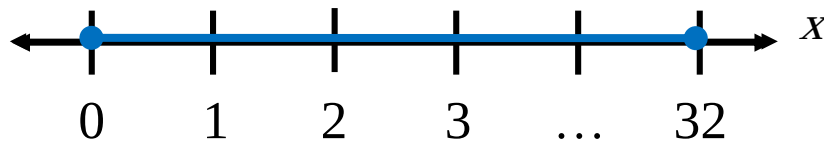
Example: Random variable

Decide whether the random variable x is discrete or continuous.

x = The volume of water in a 32-ounce container.

Solution:

Continuous random variable (The amount of water can be any volume between 0 ounces and 32 ounces)



حدد ما إذا كان المتغير العشوائي منفصلاً أم متصلاً؟
حجم الماء في وعاء سعة 32 أونصة.

الحل:

متغير عشوائي مستمر (كمية الماء يمكن أن تتراوح بين 0 أونصة و32 أونصة)

Discrete Probability Distribution

Discrete probability distribution

Lists each possible value the random variable can assume, together with its probability.

Must satisfy the following conditions:

In Words

1. The probability of each value of the discrete random variable is between 0 and 1, inclusive.
2. The sum of all the probabilities is

In Symbols

$$0 \leq P(x) \leq 1$$

$$\sum P(x) = 1$$

- التوزيعات الاحتمالات المنفصلة
يسرد كل قيمة محتملة يمكن أن يفترضها المتغير العشوائي،
بالإضافة إلى احتمالياتها
- 1- يتراوح احتمال كل قيمة للمتغير العشوائي المنفصل بين 0 و 1 ضمناً.
 - 2- مجموع كل الاحتمالات هو 1

Constructing a Discrete Probability Distribution

Let x be a discrete random variable with possible outcomes $x_1, x_2, \dots, x_n$.

x يكون متغيراً عشوائياً منفصلاً مع النتائج المحتملة $x_1, x_2, \dots, x_n$

1. Make a frequency distribution for the possible outcomes.
2. Find the sum of the frequencies.
3. Find the probability of each possible outcome by dividing its frequency by the sum of the frequencies.
4. Check that each probability is between 0 and 1 and that the sum is 1.

- 1- قم بإجراء توزيع ترددي للنتائج المحتملة.
- 2- أوجد مجموع الترددات.
- 3- أوجد احتمال كل نتيجة محتملة عن طريق قسمة تكرارها على مجموع التكرارات.
- 4- تأكد من أن كل احتمال يقع بين 0 و 1 وأن المجموع هو 1.

Example: Constructing a Discrete Probability Distribution

An industrial psychologist administered a personality inventory test for passive-aggressive traits to 150 employees. Individuals were given a score from 1 to 5, where 1 was extremely passive and 5 extremely aggressive. A score of 3 indicated neither trait. Construct a probability distribution for the random variable x .

Score, x	Frequency, f
1	24
2	33
3	42
4	30
5	21

أجرى أحد علماء النفس الصناعي اختبار جرد الشخصية للسمات العدوانية السلبية على 150 موظفًا. تم منح الأفراد درجة من 1 إلى 5، حيث كان 1 سلبيًا للغاية و5 عدوانيًا للغاية. تشير النتيجة 3 إلى عدم وجود سمة. أنشئ التوزيع الاحتمالي للمتغير العشوائي x

Solution: Constructing a Discrete Probability Distribution

Divide the frequency of each score by the total number of individuals in the study to find the probability for each value of the random variable.

$$P(1) = \frac{24}{150} = 0.16 \quad P(2) = \frac{33}{150} = 0.22 \quad P(3) = \frac{42}{150} = 0.28$$

$$P(4) = \frac{30}{150} = 0.20 \quad P(5) = \frac{21}{150} = 0.14$$

- Discrete probability distribution:

x	1	2	3	4	5
$P(x)$	0.16	0.22	0.28	0.20	0.14

Solution: Constructing a Discrete Probability Distribution

x	1	2	3	4	5
$P(x)$	0.16	0.22	0.28	0.20	0.14

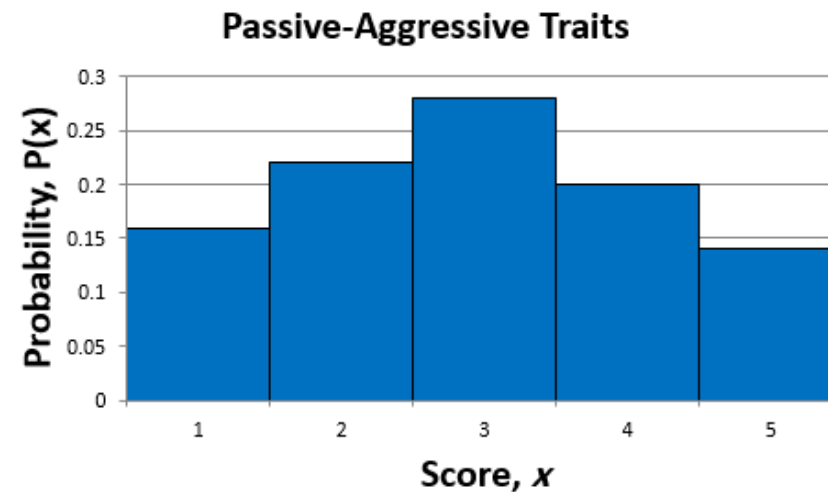


This is a valid discrete probability distribution since

1. Each probability is between 0 and 1, inclusive,
 $0 \leq P(x) \leq 1$
2. The sum of the probabilities equals 1,
 $\Sigma P(x) = 0.16 + 0.22 + 0.28 + 0.20 + 0.14 = 1$

Solution: Constructing a Discrete Probability Distribution

Histogram



Because the width of each bar is one, the area of each bar is equal to the probability of a particular outcome.

Mean, Variance and standard deviation for probability Distribution

المتوسط والتباين والانحراف المعياري للتوزيع الاحتمالي

Mean of a discrete probability distribution $\mu = \sum xP(x)$

Each value of x is multiplied by its corresponding probability and the products are added. Also, mean is known as (Expected Value).

Variance of a discrete probability distribution $\sigma^2 = \sum (x - \mu)^2 P(x)$

Standard deviation of a discrete probability distribution

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum (x - \mu)^2 P(x)}$$

Example:

The probability distribution for the personality inventory test for passive-aggressive traits is given.

Find the mean (Expected value)

Find Variance and Standard deviation

Solution:

$$\mu = \sum x P(x) = 2.94$$

x	$P(x)$	$xP(x)$
1	0.16	$1(0.16) = 0.16$
2	0.22	$2(0.22) = 0.44$
3	0.28	$3(0.28) = 0.84$
4	0.20	$4(0.20) = 0.80$
5	0.14	$5(0.14) = 0.70$

Mean, Variance and standard deviation

x	$P(x)$	$x - \mu$	$(x - \mu)^2$	$(x - \mu)^2 P(x)$
1	0.16	$1 - 2.94 = -1.94$	$(-1.94)^2 = 3.764$	$3.764(0.16) = 0.602$
2	0.22	$2 - 2.94 = -0.94$	$(-0.94)^2 = 0.884$	$0.884(0.22) = 0.194$
3	0.28	$3 - 2.94 = 0.06$	$(0.06)^2 = 0.004$	$0.004(0.28) = 0.001$
4	0.20	$4 - 2.94 = 1.06$	$(1.06)^2 = 1.124$	$1.124(0.20) = 0.225$
5	0.14	$5 - 2.94 = 2.06$	$(2.06)^2 = 4.244$	$4.244(0.14) = 0.594$

Variance

$$\sigma^2 = \Sigma(x - \mu)^2 P(x) = 1.616$$

standard deviation

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.616} \approx 1.3$$

Section 4.2 Objectives

- Determine if a probability experiment is a binomial experiment
- Find binomial probabilities using the binomial probability formula
- Find binomial probabilities using technology and a binomial table
- Graph a binomial distribution
- Find the mean, variance, and standard deviation of a binomial probability distribution

تحديد ما إذا كانت تجربة الاحتمال هي تجربة ذي الحدين
حساب الاحتمالات ذي الحدين باستخدام صيغة الاحتمال ذي الحدين
ابحث عن الاحتمالات ذي الحدين باستخدام التكنولوجيا وجدول ذي
الحدين
رسم بياني لتوزيع ذي الحدين
حساب المتوسط والتباين والانحراف المعياري للتوزيع الاحتمالي ذي
الحدين

Binomial Experiments

توزيع ذي الحدين

1. The experiment is repeated for a fixed number of trials, where each trial is independent of other trials.
2. There are only two possible outcomes of interest for each trial. The outcomes can be classified as a success (S) or as a failure (F).
3. The probability of a success $P(S)$ is the same for each trial.
4. The random variable x counts the number of successful trials.

يتم تكرار التجربة لعدد محدد من التجارب، حيث تكون كل تجربة مستقلة عن التجارب الأخرى.

لا يوجد سوى نتيجتين محتملتين مثيرتين للاهتمام لكل تجربة
نجاح S وفشل F

احتمال النجاح $P(S)$ هو نفسه لكل محاولة

المتغير العشوائي x يحسب عدد التجارب الناجحة.

Notation for Binomial Experiments

توزيع ذي الحدين

Symbol	Description
n	The number of times a trial is repeated
$p = P(s)$	The probability of success in a single trial
$q = P(F)$	The probability of failure in a single trial ($q = 1 - p$)
x	The random variable represents a count of the number of successes in n trials: $x = 0, 1, 2, 3, \dots, n.$

n عدد مرات تكرار التجربة

احتمال النجاح $P(S) = p$

احتمال الفشل $P(F) = q$

$$q = (1 - p)$$

Binomial Probability Formula

Binomial Probability Formula

The probability of exactly x successes in n trials is

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

- n = number of trials
- p = probability of success
- $q = 1 - p$ probability of failure
- x = number of successes in n trials

Example

Micro fracture knee surgery has a 75% chance of success on patients with degenerative knees. The surgery is performed on three patients. Find the probability of the surgery being successful on exactly two patients.

Solution:

$$n=3 \quad p= 75\% = 0.75 \quad q = 1- 0.75 = 0.25$$

$$C_x^n p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

تتمتع جراحة الركبة ذات الكسور الدقيقة بفرصة نجاح بنسبة 75٪ للمرضى الذين يعانون من تنكس الركبتين. يتم إجراء الجراحة على ثلاثة مرضى. أوجد احتمال نجاح العملية الجراحية لمريضين بالضبط.



Solution

$$P(x=2) = C_x^n p^x q^{n-x}$$

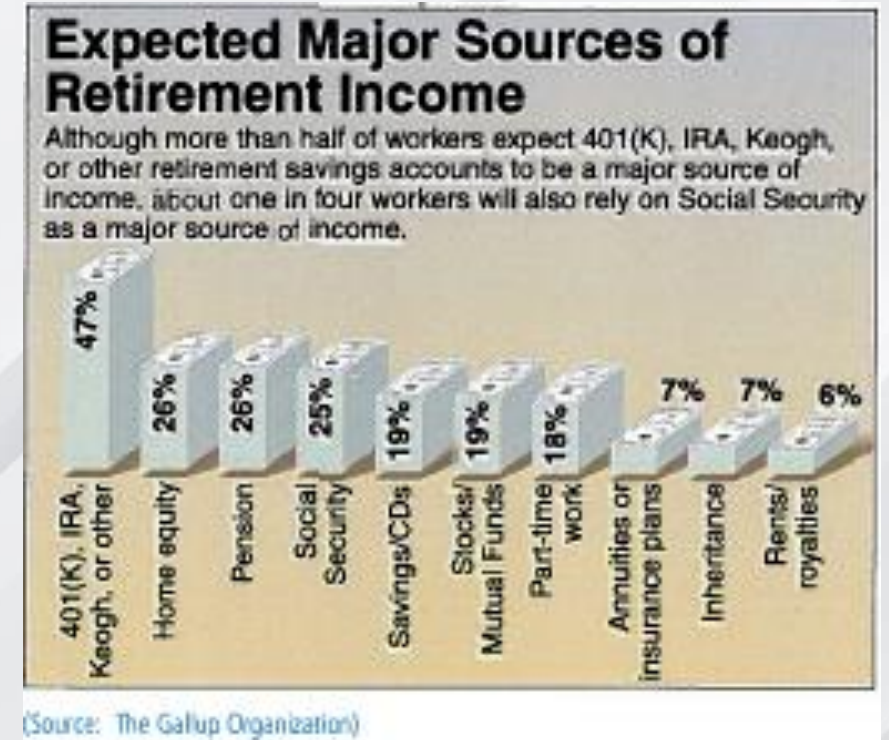
$$= C_2^3 (0.75)^2 (0.25)^{3-2}$$

$$= \frac{3!}{(3-2)! 2!} (0.75)^2 (0.25)^{3-2} = 3 (0.75)^2 (0.25)^{3-2} = 0.422$$

Example

In a survey, workers in the U.S. were asked to name their expected sources of retirement income. Seven workers who participated in the survey are randomly selected and asked whether they expect to rely on Social Security for retirement income and found it equal to 25%. Create a binomial probability distribution for the all values.

في أحد الاستطلاعات، طُلب من العاملين في الولايات المتحدة تسمية مصادرهم المتوقعة لدخل التقاعد. تم اختيار سبعة عمال شاركوا في الاستطلاع بشكل عشوائي وسئلوا عما إذا كانوا يتوقعون الاعتماد على الضمان الاجتماعي للحصول على دخل التقاعد ووجدوا أنه يساوي 25%. إنشاء توزيع احتمالي ذي الحدين لجميع القيم.



Solution

25% of working Americans expect to rely on Social Security for retirement income.

$$n = 7, \quad p = 0.25, \quad q = 0.75, \quad x = 0, 1, 2, 3, 4, 5, 6, 7$$

$$P(x=0) = {}_7C_0(0.25)^0(0.75)^7 = 1(0.25)^0(0.75)^7 \approx 0.1335$$

$$P(x=1) = {}_7C_1(0.25)^1(0.75)^6 = 7(0.25)^1(0.75)^6 \approx 0.3115$$

$$P(x=2) = {}_7C_2(0.25)^2(0.75)^5 = 21(0.25)^2(0.75)^5 \approx 0.3115$$

$$P(x=3) = {}_7C_3(0.25)^3(0.75)^4 = 35(0.25)^3(0.75)^4 \approx 0.1730$$

$$P(x=4) = {}_7C_4(0.25)^4(0.75)^3 = 35(0.25)^4(0.75)^3 \approx 0.0577$$

$$P(x=5) = {}_7C_5(0.25)^5(0.75)^2 = 21(0.25)^5(0.75)^2 \approx 0.0115$$

$$P(x=6) = {}_7C_6(0.25)^6(0.75)^1 = 7(0.25)^6(0.75)^1 \approx 0.0013$$

$$P(x=7) = {}_7C_7(0.25)^7(0.75)^0 = 1(0.25)^7(0.75)^0 \approx 0.0001$$

Example

A survey indicates that 41% of women in the U.S. consider reading their favorite leisure-time activity. You randomly select four U.S. women and ask them if reading is their favorite leisure-time activity. Find the probability that at least two of them respond yes.

تشير إحدى الدراسات إلى أن 41% من النساء في الولايات المتحدة يفكرن في القراءة كنشاط مفضل لديهن في أوقات الفراغ. قمت باختيار أربع نساء أمريكيات بشكل عشوائي واسألهن عما إذا كانت القراءة هي نشاطهن المفضل في أوقات الفراغ. أوجد احتمال أن يجيب اثنان منهم على الأقل بنعم.



Solution:

- $n = 4$, $p = 0.41$, $q = 0.59$
- At least two means two or more.
- Find the sum of $P(2)$, $P(3)$, and $P(4)$.

Solution

$$P(X=2) = {}_4C_2(0.41)^2(0.59)^2 = 6(0.41)^2(0.59)^2 \approx 0.351094$$

$$P(X=3) = {}_4C_3(0.41)^3(0.59)^1 = 4(0.41)^3(0.59)^1 \approx 0.162654$$

$$P(X=4) = {}_4C_4(0.41)^4(0.59)^0 = 1(0.41)^4(0.59)^0 \approx 0.028258$$

$$\begin{aligned} P(X \geq 2) &= P(2) + P(3) + P(4) \\ &\approx 0.351094 + 0.162654 + 0.028258 \\ &\approx 0.542 \end{aligned}$$



Mean, Variance and standard deviation for Binomial Distribution

- **Mean:** $\mu = np$
- **Variance:** $\sigma^2 = npq$
- **Standard Deviation:** $\sigma = \sqrt{npq}$

Example

In Pittsburgh, Pennsylvania, about 56% of the days in a year are cloudy. Find the mean, variance, and standard deviation for the number of cloudy days during the month of June. (*Source: National Climatic Data Center*)



Solution: $n = 30$, $p = 0.56$, $q = 0.44$

Mean: $\mu = np = 30 \cdot 0.56 = 16.8$

Variance: $\sigma^2 = npq = 30 \cdot 0.56 \cdot 0.44 \approx 7.4$

Standard Deviation:

في بيتسبرغ، بنسلفانيا، حوالي 56% من أيام السنة غائمة.
أوجد المتوسط والتباين والانحراف المعياري لعدد الأيام
الغائمة خلال شهر يونيو. (المصدر: المركز الوطني
للبيانات المناخية)

4.3 More Discrete Probability Distributions

- Find probabilities using the geometric distribution
أوجد الاحتمالات باستخدام التوزيع الهندسي
- Find probabilities using the Poisson distribution
إيجاد الاحتمالات باستخدام توزيع بواسون

Geometric distribution

Geometric distribution

A discrete probability distribution.

Satisfies the following conditions:

- 1- A trial is repeated until a success occurs.
- 2- The repeated trials are independent of each other.
- 3- The probability of success p is constant for each trial.
- 4- The probability that the first success will occur on trial x is

$$P(x) = p(q)^{x-1}, \quad \text{where } q = 1 - p.$$

تكرر التجربة حتى تنجح.
التجارب المتكررة مستقلة عن بعضها البعض.
احتمال النجاح ثابت لكل تجربة.
احتمال أن يحدث النجاح الأول في التجربة هو x

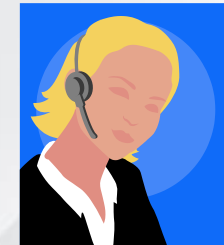
Example

From experience, you know that the probability that you will make a sale on any given telephone call is 0.23. Find the probability that your first sale on any given day will occur on your fourth or fifth sales call.

من خلال تجربتك، تعلم أن احتمال إجراء عملية بيع في أي مكالمات هاتفية هو 0.23. ابحث عن احتمالية أن تتم عملية البيع الأولى في أي يوم في مكالمات المبيعات الرابعة أو الخامسة.

Solution:

- $P(\text{sale on fourth or fifth call}) = P(4) + P(5)$
- Geometric with $p = 0.23$, $q = 0.77$, $x = 4, 5$



Solution

- $P(4) = 0.23(0.77)^{4-1} \approx 0.105003$
- $P(5) = 0.23(0.77)^{5-1} \approx 0.080852$

$$\begin{aligned} P(\text{sale on fourth or fifth call}) &= P(4) + P(5) \\ &\approx 0.105003 + 0.080852 \\ &\approx 0.186 \end{aligned}$$



Poisson distribution

A discrete probability distribution.

Satisfies the following conditions:

- The experiment consists of counting the number of times an event, x , occurs in a given interval.
- The interval can be an interval of time, area, or volume.
- The probability of the event occurring is the same for each interval.
- The number of occurrences in one interval is independent of the number of occurrences in other intervals.

تتكون التجربة من حساب عدد مرات حدوث في الحدث X فترة زمنية معينة.

يمكن أن يكون الفاصل الزمني فاصلًا زمنيًا أو مساحة أو حجمًا.

احتمال وقوع الحدث هو نفسه لكل فترة.

عدد التكرارات في فترة زمنية واحدة مستقل عن عدد التكرارات في فترات زمنية أخرى.

Poisson distribution

Conditions continued:

The probability of the event occurring is the same for each interval.

The probability of exactly x occurrences in an interval is

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

where $e \approx 2.71818$ and μ is the mean number of occurrences

تابع الشروط:

احتمال وقوع الحدث هو نفسه لكل فترة.

احتمال حدوث الحدث بالضبط في فترة ما هو يعطى بالقانون

Example

The mean number of accidents per month at a certain intersection is 3. What is the probability that in any given month four accidents will occur at this intersection?



متوسط عدد الحوادث شهريًا عند تقاطع معين هو 3. ما احتمال وقوع أربع حوادث في أي شهر عند هذا التقاطع؟

Solution

Solution:

- Poisson with $x = 4, \mu = 3$

$$P(4) = \frac{3^4(2.71828)^{-3}}{4!} \approx 0.168$$

Un Solved Examples

Chapter

If in Binomial distribution $n = 32$, $p = 0.7$, $q = 0.3$ then Find mean μ and
Variance σ^2 and Standard Deviation σ

Un Solved Examples

Chapter

If in Geometric distribution $p = 0.2$, $q = 0.8$ then Find $P(4)$

Un Solved Examples

Chapter

Number of rooms in the hospital is a discrete random variable

Chapter Summary

Mean of a discrete probability distribution = Expected value = $\mu = \sum x P(x)$

Variance of a discrete probability distribution $\sigma^2 = \sum (x - \mu)^2 P(x)$

Standard deviation of a discrete probability distribution

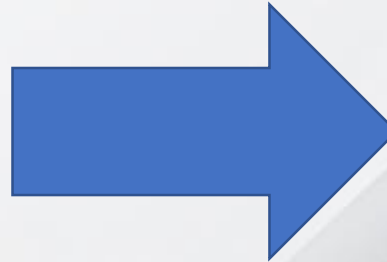
Chapter Summary

Chapter

Binomial Distributions :

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

Where; $q = 1-p$



Binomial Distributions :

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

Where; $q = 1-p$

Chapter Summary

Chapter

Geometric Distribution:

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

Poisson Distribution:

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



Thank You