**1. Пространство элементарных событий. Классификация случайных событий. Алгебра событий.**

**Пространство элементарных событий** (Ω) представляет собой множество всех возможных исходов эксперимента. Элементы этого множества называются элементарными событиями.

**Случайное событие** - любое подмножество пространства элементарных событий. То есть случайное событие A - набор определенных исходов эксперимента.

**Классификация случайных событий:**

1. Достоверное событие (Ω): Событие, которое происходит всегда, в любом исходе эксперимента.

2. Невозможное событие (∅): Событие, которое не происходит никогда.

3. Противоположное событие (A'): Все исходы, не входящие в событие A.

4. Совместные события (A ∩ B): Событие, которое происходит, если происходят оба события A и B.

5. Объединение событий (A ∪ B): Событие, которое происходит, если происходит хотя бы одно из событий A или B.

6. Разностные события (A B): Событие, которое происходит, если происходит событие A, но не происходит B.

**Алгебра событий** включает в себя операции над событиями, аналогичные операциям над множествами. Например, пересечение (A ∩ B), объединение (A ∪ B), дополнение (A'), разность (A B) и так далее.

Если у нас есть два события A и B, то алгебра событий позволяет нам определить различные комбинации этих событий и рассчитать вероятность их происхождения. При этом важно использовать правила вероятности, такие как правило сложения и правило умножения, чтобы корректно оценивать вероятности совместных и независимых событий.

**2. Статистическое определение вероятности**

**Статистическое определение вероятности** основано на представлении вероятности как относительной частоты появления определенного события в серии экспериментов. Формально, оно выражается следующим образом:

Пусть у нас есть эксперимент, который может быть повторен много раз с одинаковыми условиями, и пусть N - общее число повторений этого эксперимента. Пусть n - число раз, когда интересующее нас событие произошло в течение этих N повторений.

Тогда вероятность события A (P(A)) определяется как предел отношения числа благоприятных исходов к общему числу исходов при бесконечном числе повторений эксперимента:



Это определение вероятности основывается на идее, что вероятность события можно приблизить, проводя большое количество независимых экспериментов, и оценивая отношение числа благоприятных исходов к общему числу исходов.

Важно отметить, что при конечном числе экспериментов эта оценка может быть не очень точной, но с увеличением числа повторений эксперимента вероятность события стремится к теоретической вероятности, которую мы определяем статистическим методом.

**3. Элементы комбинаторики: размещения, сочетания, перестановки**

**Элементы комбинаторики** - это методы подсчета и оценки числа различных комбинаций и распределений элементов в различных ситуациях. Три основных элемента комбинаторики, о которых вы упомянули, включают в себя размещения, сочетания и перестановки. Давайте рассмотрим каждый из них более подробно:

1. **Размещения** (перестановки с повторениями):

Размещения представляют собой упорядоченные упорядоченные выборки из элементов множества. Если у нас есть множество из n элементов, и мы выбираем k элементов для размещения, то количество размещений обозначается как(A\_n^k\) и вычисляется следующим образом:



2. **Сочетания** (перестановки без повторений):

Сочетания представляют собой неупорядоченные выборки из элементов множества. Если у нас есть множество из n элементов, и мы выбираем k элементов для сочетания, то количество сочетаний обозначается как(C\_n^k\) и вычисляется следующим образом:



3. **Перестановки** (перестановки без повторений):

Перестановки представляют собой упорядоченные упорядоченные выборки из элементов множества. Если у нас есть множество из n элементов, то количество перестановок обозначается как(P\_n\) и вычисляется следующим образом:

P\_n = n!

Эти элементы комбинаторики широко используются в теории вероятностей, статистике и других областях для решения задач подсчета и анализа возможных комбинаций элементов.

**4. Классическое определение вероятности**

Классическое определение вероятности основано на предположении, что все элементарные исходы эксперимента равновозможны и что каждый из них имеет одинаковую вероятность произойти. Это определение часто применяется в ситуациях, где мы имеем дело с равновозможными исходами.

Формально, если у нас есть событие A, которое может произойти в одном из n равновозможных исходов, то вероятность события A (P(A)) вычисляется как:



Или если обозначить через( N(A)) количество благоприятных исходов для события A, а через( N) общее количество возможных исходов в эксперименте:



Важно отметить, что это определение применимо только в тех ситуациях, где все исходы равновозможны и их количество конечно. Например, бросание симметричного кубика, где у нас есть 6 равновозможных исходов (выпадение одной из шести граней).

Классическое определение вероятности часто используется в теории вероятностей, особенно когда мы имеем дело с простыми и равновозможными случаями.

**5. Геометрическое определение вероятности**

Геометрическое определение вероятности связано с измерением геометрических фигур и областей на пространстве элементарных событий. Это определение часто используется в ситуациях, когда события можно представить в виде геометрических фигур или областей.

Предположим, что у нас есть некоторая область S, представляющая собой пространство элементарных событий. Пусть A будет событием, которое можно представить как геометрическую фигуру внутри области S. Тогда вероятность события A, обозначаемая как P(A), определяется как отношение площади (или объема в трехмерных пространствах) геометрической фигуры A к общей площади (или объему) пространства элементарных событий S.

Математически это выражается следующим образом:



Это определение вероятности часто применяется в задачах, связанных с геометрическими вероятностями, например, когда мы бросаем точку на плоскость или вводим случайную величину с равномерным распределением в некотором геометрическом пространстве.

**6. Аксиоматическое построение теории вероятностей**

При аксиоматическом построении теории вероятностей первичным понятием является не элементарное случайное событие, а просто элементарное событие любой природы. Множество таких событий образует поле элементарных событий. Из подмножества данного множества составляются некоторые ансамбли, которые и носят название случайного события. Множество таких событий образует поле событий *S.* На этом поле случайных событий вводится числовая функция, называемая вероятностью и определяемая следующими аксиомами.

Аксиоматический подход к построению теории вероятностей основывается на нескольких аксиомах или постулатах о поведении вероятности. Рассмотрим основные аксиомы:

• **Ненегативность**. Вероятности всех событий неотрицательны:

P(A) ≥ 0 для любого события A.

• **Нормировка**. Вероятность достоверного события равна 1:

P(S) = 1, где S - пространство элементарных событий.

• **Аддитивность**. Если A и B - несовместные события, то P(A∪B) = P(A) + P(B). Иными словами, вероятность объединения несовместных событий равна сумме вероятностей.

Из этих аксиом выводятся все основные теоретические положения теории вероятностей: теоремы о вероятности противоположного события, вероятности объединения событий, условной вероятности и т.д.

Такой аксиоматический подход придаёт теории вероятностей строгость, логичность и непротиворечивость. Однако при этом требуется дополнительное обоснование согласованности аксиом с реальными случайными экспериментами и наблюдениями.

**7. Условная вероятность. Теорема умножения вероятностей.**

**Условная вероятность** - это вероятность события A при условии, что произошло событие B, и обозначается как P(A | B). Формула для условной вероятности выражается следующим образом:



где P(A ∩ B) - вероятность одновременного наступления событий A и B, а P(B) - вероятность события B.

**Теорема умножения вероятностей** обеспечивает способ вычислить вероятность одновременного наступления двух событий. Для двух событий A и B формула теоремы умножения записывается как:

P(A∩B): Это вероятность одновременного наступления событий A и B, т.е., вероятность того, что произойдут оба события A и B.

P(A∣B): Это условная вероятность события A при условии, что событие B произошло.



Эти формулы позволяют вычислить вероятность одновременного наступления событий A и B через условные вероятности.

**8. Формула полной вероятности**

Формула полной вероятности используется для вычисления вероятности события A, разбитого на несколько случаев (или гипотез), учитывая все возможные исходы. Если {B₁, B₂, ..., Bn} - полная система событий (то есть эти события покрывают все возможные исходы и являются попарно несовместными), то формула полной вероятности записывается следующим образом:



Где:

- P(A | B\_i) - условная вероятность события A при условии, что событие Bᵢ произошло.

- P(B\_i) - вероятность того, что произойдет событие Bᵢ.

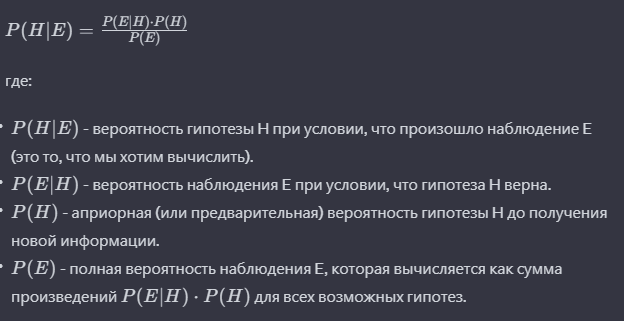
Эта формула предполагает, что события {B₁, B₂, ..., Bn} формируют полную систему событий, то есть каждый возможный исход включен в одно из событий Bᵢ, и эти события попарно несовместны.

Формула полной вероятности особенно полезна, когда происходит разбиение пространства элементарных событий на непересекающиеся категории, и мы хотим вычислить вероятность события A в зависимости от этих категорий.

**9. Формула Байеса**

Формула Байеса (или теорема Байеса) является одним из фундаментальных результатов теории вероятностей и статистики. Она используется для вычисления условных вероятностей.

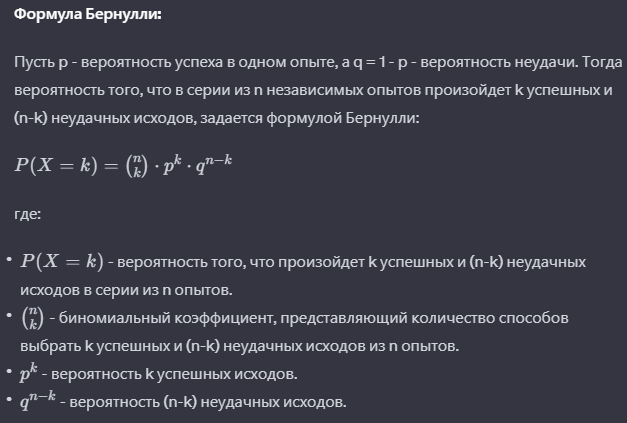
Формула Байеса устанавливает связь между безусловной и условной вероятностью следующим образом:



Эта формула полезна для обновления вероятностных оценок гипотез, учитывая новые данные

**10. Повторение независимых опытов. Формула Бернулли**

Повторение независимых опытов - это ситуация, когда один и тот же эксперимент проводится несколько раз, и каждый раз результат не зависит от предыдущих исходов. Такие опыты часто описываются формулой Бернулли, которая предоставляет вероятность наступления определенного количества успешных (или неуспешных) исходов в серии независимых испытаний.



**11.Закон распределения дискретной случайной величины**

Закон распределения дискретной случайной величины описывает вероятности возможных значений этой величины. Для дискретной случайной величины X, закон распределения определяется как функция P(X), которая определяет вероятность того, что X примет определенное значение.

Закон распределения дискретной случайной величины может быть представлен в виде таблицы или формулы, где указываются все возможные значения X и соответствующие вероятности.

Примером дискретной случайной величины является бросок кубика, где X - число выпавших очков. Закон распределения для этого случая будет следующим:

P(1) = p1

P(2) = p2

P(3) = p3

P(4) = p4

P(5) = p5

P(6) = p6

Где p1+p2+p3+p4+p5+p6 = 1 и каждое p – вероятность опрелелённого события.

**12.Функция распределения случайной величины и её свойства**

Функция распределения F(x) случайной величины X задаётся следующим образом:

F(x) = P(X ≤ x) для любого значения x из области определения X.

Иными словами, функция распределения показывает вероятность того, что случайная величина X примет значение, меньшее либо равное данному порогу x.

Основные свойства функции распределения:

• 0 ≤ F(x) ≤ 1

• F(x) не убывает

• lim F(x) = 0, lim F(x) = 1

x→-∞ x→+∞

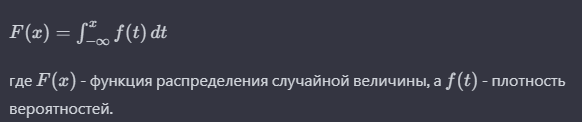
• Если x1 < x2, то F(x2) - F(x1) = P(x1 < X ≤ x2)

Кроме того, для непрерывных случайных величин функция распределения дифференцируема и её производная равна плотности вероятности f(x). А для дискретных случайных величин F(x) является кусочно-постоянной функцией.

Таким образом, зная вид функции распределения, можно полностью описать поведение случайной величины.

**13.Плотность вероятностей и её свойства**

Плотность вероятности f(x) используется для описания непрерывных случайных величин. Она связана с функцией распределения F(x) следующим соотношением:

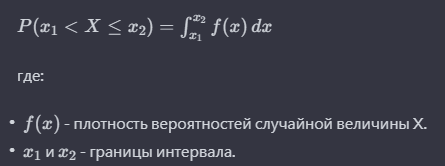


**Основные свойства плотности вероятности:**

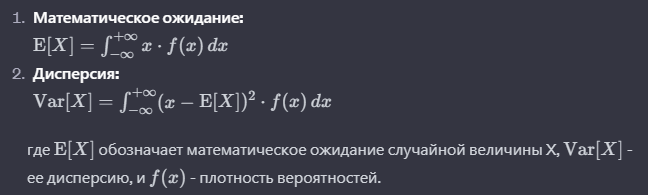
• f(x) >= 0 при всех значениях x (неотрицательность)

• Интеграл от f(x) по всей области определения равен 1 (нормировка)

• Площадь под кривой плотности вероятности между x1 и x2 равна вероятности попадания X в этот интервал:



• Для нахождения наиболее вероятного значения X нужно найти максимум f(x)

• 

Таким образом, зная вид плотности распределения, можно полностью описать поведение непрерывной случайной величины.

**14.Числовые характеристики положения случайной величины.**

Числовые характеристики положения случайной величины описывают её положение или центральную тенденцию. Некоторые из основных числовых характеристик положения включают:

**1. Математическое ожидание (ожидаемое значение):**

Математическое ожидание случайной величиныX, обозначаемое какM(X), представляет собой среднее значение или центральную точку распределения. Для дискретной случайной величиныX:



А для непрерывной случайной величины, интеграл заменяет сумму.

**2. Среднеквадратическое отклонение:**

Среднеквадратическое отклонение (или стандартное отклонение) σ измеряет разброс значений случайной величины относительно её математического ожидания.

σ = sqrt(D)

3. Дисперсия:

Дисперсия (D(X) = M[(X - M(X))^2]) — это среднее значение квадратов отклонений случайной величины от её математического ожидания.

Эти числовые характеристики положения помогают описывать форму и особенности распределения случайных величин. Выбор конкретной характеристики может зависеть от свойств данных и целей анализа.

**15.Числовые характеристики рассеивания случайной величины**

Основные числовые характеристики, описывающие рассеивание (разброс) случайной величины:

1. Дисперсия. Характеризует средний квадрат отклонения случайной величины от её математического ожидания:

D(X) = M[(X - M(X))^2]

2. Среднее квадратичное отклонение. Квадратный корень из дисперсии. Измеряется в тех же единицах, что и сама случайная величина:

σ(X) = √D(X)

3. Коэффициент вариации. Отношение среднего квадратичного отклонения к математическому ожиданию. Безразмерная величина, позволяющая сравнивать рассеивание разных случайных величин:

CV(X) = σ(X) / M(X)

4. Размах вариации. Разность между наибольшим и наименьшим значениями случайной величины.

В целом чем больше показатели рассеивания, тем сильнее отдельные значения случайной величины отличаются от её среднего уровня.

**16.Моменты случайных величин. Асимметрия и эксцесс**

Моменты случайных величин - это числовые характеристики распределения вероятностей случайной величины. Четыре основных момента - это математическое ожидание, дисперсия, асимметрия и эксцесс.

**• Математическое ожидание (первый момент):**

Математическое ожидание случайной величины X обозначается как

M(X) и представляет собой среднее значение случайной величины. Для дискретной случайной величины это сумма произведений значений на их вероятности, а для непрерывной - интеграл.



• **Дисперсия (второй момент):**

Дисперсия случайной величины X обозначается как D(X) и измеряет разброс значений вокруг среднего.



• **Асимметрия (третий момент):**

Асимметрия характеризует степень и направление отклонения формы распределения от симметричной. Если распределение смещено вправо, то асимметрия положительная, если влево - отрицательная.



• **Эксцесс (четвертый момент):**

Эксцесс измеряет степень остроты или плоскости пика распределения случайной величины. Для нормального распределения эксцесс равен 3



**17.Характеристическая функция**

Характеристическая функция случайной величины - функция, полностью описывающая распределение этой случайной величины. Она представляет собой математическое ожидание экспоненциальной функции комплексного аргумента, связанного с распределением случайной величины.

Характеристическая функция ϕ(t) случайной величины Х определяется следующим образом:

ϕ(t)=E(e ^ itX)

где:

- t - параметр, который представляет собой комплексное число.

- i - мнимая единица (i^2 = -1).

- Х- случайная величина.

Эта функция играет важную роль в теории вероятностей и статистике по нескольким причинам:

1. Уникальность распознавания:

Характеристическая функция однозначно определяет распределение случайной величины. Для двух случайных величин с одинаковыми характеристическими функциями существует однозначное соответствие между их распределениями.

2. Вычисление моментов:

Моменты случайной величины могут быть выражены через производные характеристической функции. Например,(n)-й момент можно найти, взяв(n)-ю производную характеристической функции в нуле.

3. Центральная предельная теорема:

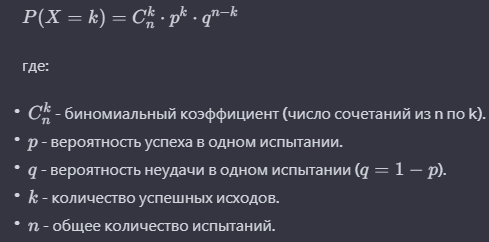
Центральная предельная теорема утверждает, что сумма большого числа независимых и одинаково распределенных случайных величин имеет распределение, близкое к нормальному.

Форма характеристической функции может существенно различаться в зависимости от характера распределения случайной величины. Например, для нормального распределения, характеристическая функция имеет вид, аналогичный экспоненциальной функции.

**18.Биномиальный закон распределения и его числовые характеристики**

**Биномиальное распределение** описывает вероятность k успешных исходов в серии из n независимых однотипных испытаний, где каждый исход имеет одинаковую вероятность успеха p и вероятность неудачи q = 1 - p.

Функция вероятности биномиального распределения P(X=k) задается формулой:



Числовые характеристики биномиального распределения:

1. **Математическое ожидание (среднее):**

Математическое ожидание биномиального распределения вычисляется по формуле:

M(X) = np

**2. Дисперсия:**

Дисперсия биномиального распределения вычисляется по формуле:

D(X) = npq

**3. Стандартное отклонение:**

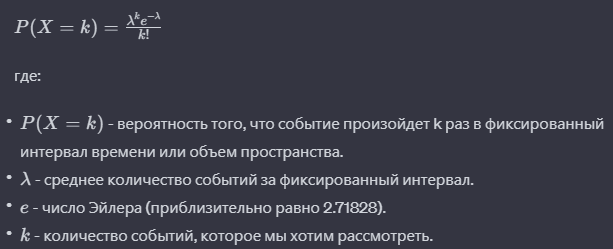
Стандартное отклонение sigma\)) равно квадратному корню из дисперсии:

σ= sqrt(npq)

Биномиальное распределение широко используется для моделирования ситуаций, где проводится фиксированное число испытаний, и каждое из них может закончиться успехом или неудачей.

**19.Закон распределения Пуассона**

Закон распределения Пуассона описывает вероятность того, что событие произойдет определенное количество раз в фиксированный интервал времени или в фиксированном объеме пространства. Этот закон применяется, когда события происходят независимо друг от друга, и вероятность каждого события постоянна.



Числовые характеристики распределения Пуассона:

λ = np

1. Математическое ожидание (среднее):

Математическое ожидание распределения Пуассона равно параметру(lambda):

M(X) = λ

2. Дисперсия:

Дисперсия распределения Пуассона также равна параметру λ

D = λ

3. Стандартное отклонение:

Стандартное отклонение sigma\)) равно квадратному корню из дисперсии:

σ=sqrt(λ)

Распределение Пуассона задаётся одним параметром λ и позволяет моделировать редкие случайные явления

**20.Простейший поток событий**

**Простейший поток событий** - вероятностная модель, описывающая процесс появления однородных случайных событий во времени.

Основные свойства простейшего потока:

• События происходят по одному, не группируясь.

• Число событий в любых неперекрывающихся интервалах времени независимо.

• Вероятность наступления события в бесконечно малом интервале пропорциональна его длине и не зависит от положения интервала на оси времени.

• Математическое ожидание числа событий прямо пропорционально длине рассматриваемого интервала времени.

Простейший поток событий используется для моделирования случайных процессов во времени, например, потока заявок, радиоактивного распада и др.

**21.Равномерный закон распределения**

**Равномерное распределение** - тип вероятностного распределения, при котором все значения внутри определенного интервала имеют одинаковую вероятность появления. Другими словами, вероятность того, что случайная величина примет любое значение в этом интервале, равномерно распределена.

Плотность вероятности для равномерного распределения:

Функция плотности вероятности для равномерного распределения на интервале([a, b]) определяется следующим образом:

f(x) =1/b - a, a<=x<=b

где:

-( a) и( b) - концы интервала.

- f(x) - значение плотности вероятности в точке x.

Функция распределения для равномерного распределения:

Функция распределения( F(x)) для равномерного распределения на интервале([a, b]\) задается как:

F(x) =x - a/b – a, a<=x<=b

Основные характеристики равномерного распределения:

• Математическое ожидание (среднее):

Математическое ожидание для равномерного распределения вычисляется как среднее значение концов интервала:

M = a+b/2

• Дисперсия:

Дисперсия равномерного распределения рассчитывается по формуле:

D = (b-a)^2/12

• Стандартное отклонение:

Стандартное отклонение σ равно квадратному корню из дисперсии:

σ = b-a/sqrt(12)

Равномерное распределение широко используется в статистике и вероятностной теории в различных областях, например, при моделировании случайных величин в определенном диапазоне значений.

**22.Показательный закон распределения**

Показательное (экспоненциальное) распределение - тип вероятностного распределения, описывающего время между последовательными независимыми событиями, которые происходят с постоянной интенсивностью. Это распределение часто используется для моделирования времени между приходами в системы обслуживания, обработки запросов, выхода из строя технических устройств и т. д.

**Функция плотности вероятности** f(x) для показательного распределения с параметром λ выражается: f(x) = λe ^ -λx, x>= 0

где:

- λ - параметр интенсивности (или обратной длительности времени между событиями).

- x - случайная величина, представляющая время между событиями.

**Функция распределения** F(x) для показательного распределения с параметром λ выглядит следующим образом:

F(x) = 1 - e^-λ x, x>=0

Основные характеристики показательного распределения:

• Математическое ожидание (среднее):

Математическое ожидание для показательного распределения равно обратной интенсивности:

M = 1/λ

• Дисперсия:

Дисперсия показательного распределения равна квадрату обратной интенсивности:

D = 1/λ^2

• Стандартное отклонение:

Стандартное отклонение (\(\sigma\)) равно обратной интенсивности:

σ =1/λ

Показательное распределение имеет отсутствие памяти, что означает, что вероятность того, что событие произойдет в следующий момент времени, не зависит от того, сколько времени прошло с момента предыдущего события. Это свойство часто используется в моделировании случайных процессов, где события происходят независимо.

**23.Нормальный закон распределения (закон Гаусса)**

Нормальное распределение, также известное как распределение Гаусса или закон Гаусса, является одним из самых важных и широко используемых распределений в статистике. Оно описывает множество случайных переменных в природе, обществе и науке. Нормальное распределение имеет колоколообразную форму и характеризуется несколькими ключевыми свойствами.

Функция плотности вероятности (PDF) для нормального распределения с параметрами среднего μ и стандартного отклонения σ задается формулой:

f(x) = (1/ σ sqrt(2 π) )\* e ^ - (x- μ)^2/2 σ^2

где:

- μ - среднее (математическое ожидание) распределения.

- σ - стандартное отклонение распределения.

- π - число Пи (приблизительно равно 3.14159).

- e - число Эйлера (приблизительно равно 2.71828).

Характеристики нормального распределения:

1. Симметрия:

Нормальное распределение симметрично относительно своего среднего значения.

2. Математическое ожидание (среднее):

Математическое ожидание нормального распределения равно параметру среднего:

M = μ

3. Дисперсия:

Дисперсия нормального распределения равна квадрату стандартного отклонения:

D = σ^2

4. Стандартное отклонение:

Стандартное отклонение (\(\sigma\)) задает ширину и разброс распределения.

5. Форма колокола:

Нормальное распределение имеет колоколообразную форму, и большинство значений сосредоточены вблизи среднего значения.

6. Правило 68-95-99.7:

Приблизительно 68% значений находятся в пределах одного стандартного отклонения от среднего, 95% - в пределах двух стандартных отклонений, и 99.7% - в пределах трех стандартных отклонений.

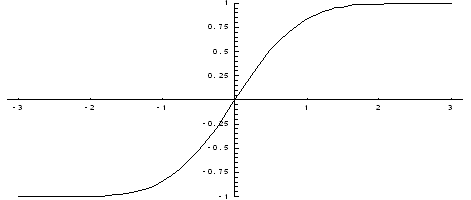
Нормальное распределение широко используется в статистике, эконометрике, физике, биологии и других областях, часто используется в статистическом анализе и моделировании данных.

**24.Функция Лапласа и ее свойства**



 Значения этой функции при различных значениях *х* посчитаны и приводятся в специальных таблицах.

            Ниже показан график функции Лапласа.



Свойства:

Ф(0) = 0 - значение функции Лапласа в нуле равно нулю.

Ф(inf) = 1 - значение функции Лапласа в бесконечности равно единице

Ф(-х) = Ф(х) - функция Лапласа от отражения функции относительно оси времени равна отрицанию функции Лапласа исходной функции.

**25.Правило трех сигм**

Правило трех сигм, также известное как правило 68-95-99.7, является статистическим правилом, определяющим, какие процентные доли значений распределены вокруг среднего в нормальном распределении. Это правило основано на нормальном распределении и указывает, какие доли данных лежат внутри различных интервалов вокруг среднего значения.

Согласно правилу трех сигм:

1. 68% данных:

Приблизительно 68% значений лежат в пределах одного стандартного отклонения от среднего значения.

2. 95% данных:

Приблизительно 95% значений лежат в пределах двух стандартных отклонений от среднего значения.

3. 99.7% данных:

Приблизительно 99.7% значений лежат в пределах трех стандартных отклонений от среднего значения.

Это правило иллюстрирует, как распределены значения в стандартном нормальном распределении (нормальное распределение с средним μ = 0 и стандартным отклонением σ = 1 и может быть применено к любому нормальному распределению с произвольными параметрами.

Применение правила трех сигм позволяет оценить, насколько значения отклоняются от среднего и дает представление о разбросе данных. Однако, важно помнить, что это правило основано на предположении нормального распределения данных, и в реальных ситуациях оно может быть менее точным для не нормально распределенных данных.

**26.Системы случайных величин**

Система случайных величин – это совокупность двух или более случайных величин, определенных на одном вероятностном пространстве.

Основные понятия:

1. Закон распределения – задает совместное распределение вероятностей для всех случайных величин системы.
2. Маргинальное распределение отдельно взятой случайной величины из системы.
3. Независимость – случайные величины называются независимыми, если любая из них не зависит от совокупных значений остальных.
4. Корреляционный момент – показывает степень линейной зависимости между случайными величинами

Анализ систем позволяет комплексно описывать сложные вероятностные эксперименты и процессы.

**27.Закон распределения дискретной системы двух случайных величин**

**28.Функция распределения системы двух случайных величин**

Функция распределения системы двух случайных величин (X, Y) - это функция, которая определяет вероятность того, что случайная величина (X, Y) примет значения (x, y). Она обозначается как F(x, y), и ее определение следующее: **F(x, y) = P(X ≤ x, Y ≤ y)**

функция распределения представляет собой таблицу, в которой каждому набору значений (x, y) соответствует вероятность того, что случайная величина (X, Y) примет эти значения.

В непрерывном случае функция распределения представляет собой непрерывную функцию, которая принимает значения от 0 до 1.

Функция распределения системы двух случайных величин имеет следующие свойства:

F(x, y) ≥ 0 для всех x и y.

F(-∞, -∞) = 0.

F(∞, ∞) = 1.

F(x, -∞) = P(X ≤ x)

F(-∞, y) = P(Y ≤ y)

F(x, y) = F(x, -∞) + F(-∞, y) - F(-∞, -∞) = P(X ≤ x) + P(Y ≤ y) - 1

**29.Плотность вероятностей системы двух случайных величин**

Плотность вероятностей системы двух случайных величин (X, Y) - это функция, которая определяет вероятность того, что случайная величина (X, Y) примет значения (x, y) в окрестности точки (x, y). Она обозначается как f(x, y), и ее определение следующее:

**f(x, y) = lim Δx→0 lim Δy→0 [F(x + Δx, y + Δy) - F(x, y)] / (Δx Δy)**

где F(x, y) - функция распределения системы двух случайных величин.

В дискретном случае плотность распределения представляет собой сумму вероятностей того, что случайная величина (X, Y) примет значения (x, y) или (x + Δx, y) или (x, y + Δy) или (x + Δx, y + Δy).

В непрерывном случае плотность распределения представляет собой непрерывную функцию, которая принимает значения в окрестности точки (x, y).

Плотность распределения системы двух случайных величин имеет следующие свойства:

f(x, y) ≥ 0 для всех x и y.

∫∫ f(x, y) dx dy = 1

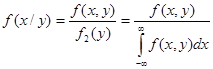
Плотность распределения системы двух случайных величин используется для вычисления вероятностей различных событий, связанных с двумя случайными величинами.

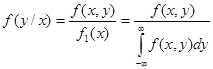
**30.Условные законы распределения случайных величин.**

Распределение одной случайной величины, входящей в систему, найденное при условии, что другая случайная величина приняла определенное значение, называется **Условным законом распределения**.

Условный закон распределения можно задавать как функцией распределения так и плотностью распределения.

Условная плотность распределения вычисляется по формулам





**31.Основные числовые характеристики системы двух случайных величин**

Основные числовые характеристики системы двух случайных величин X и Y - это характеристики, которые описывают основные свойства этих величин. К основным числовым характеристикам относятся:

*Математическое ожидание.* Математическое ожидание случайной величины X обозначается как μX и определяется следующим образом:

μX = E(X) = ∫xf(x)dx

где f(x) - плотность распределения случайной величины X.

Математическое ожидание характеризует среднее значение случайной величины.

*Дисперсия*. Дисперсия случайной величины X обозначается как σ^2X и определяется следующим образом:

σ^2X = E((X - μX)^2) = ∫(x - μX)^2f(x)dx

Дисперсия характеризует разброс значений случайной величины относительно ее математического ожидания.

*Среднеквадратическое отклонение*. Среднеквадратическое отклонение случайной величины X обозначается как σX и определяется следующим образом:

σX = √σ^2X

Среднеквадратическое отклонение характеризует степень разброса значений случайной величины относительно ее математического ожидания.

Кроме того, для системы двух случайных величин X и Y используются следующие числовые характеристики:

*Ковариация*. Ковариация случайной величины X и случайной величины Y обозначается как Cov(X, Y) и определяется следующим образом:

Cov(X, Y) = E((X - μX)(Y - μY)) = ∫(x - μX)(y - μY)f(x, y)dx dy

Ковариация характеризует взаимосвязь между двумя случайными величинами.

*Корреляция*. Корреляция между случайной величиной X и случайной величиной Y обозначается как ρ(X, Y) и определяется следующим образом:

ρ(X, Y) = Cov(X, Y) / (σXσY)

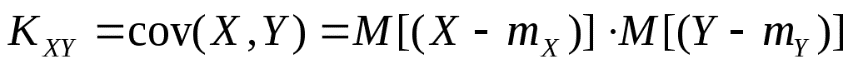
Корреляция характеризует степень линейной зависимости между двумя случайными величинами.

Числовые характеристики системы двух случайных величин используются для анализа различных свойств этих величин. Например, математическое ожидание и дисперсия используются для оценки среднего значения и разброса значений случайной величины. Ковариация и корреляция используются для оценки взаимосвязи между двумя случайными величинами.

**32.Корреляционный момент. Коэффициент корреляции и его свойства**

Для описания системы двух случайных величин, кроме математических ожиданий и дисперсий составляющих пользуются и другими характеристиками, к числу которых относятся *корреляционный момент*и*коэффициент корреляции*

*Корреляционным моментом*(или*ковариацией,*или*моментом связи*) двух случайных величин *X*и*Y*называется произведение отклонений этих величин



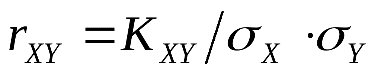
Корреляционный момент служит для характеристики связи между величинами *X*и*Y.*

Корреляционный момент равен нулю, если *X*и*Y*являютсянезависимыми;

*Следовательно, если корреляционный момент не равен нулю, то X и Y – зависимые случайные величины.*

Величина корреляционного момента зависит от единиц измерения случайных величин. По этой причине для одних и тех же двух величин*X*и*Y,*величина корреляционного момента будет иметь различные значения в зависимости от того, в каких единицах были измерены величины.

Для того чтобы устранить этот недостаток, вводят новую числовую характеристику- - «*коэффициент корреляции*».

*Коэффициентом корреляции *случайных величин**иназывают отношение корреляционного момента к произведению средних квадратических отклонений этих величин: 

R – безразмерная величина, что удобнее

**Свойства коэффициента корреляции:**

1. Коэффициент корреляции по абсолютной величине непревосходит 1

2. Если случайные величины**инезависимы, токоэффициент корреляции равен нулю

**33.Условные математические ожидания. Линии регрессии**

Условное математическое ожидание - это математическое ожидание случайной величины, вычисленное при условии, что другая случайная величина приняла некоторое значение.

Например, рассмотрим две случайные величины: рост человека (Y) и его вес (X). Условное математическое ожидание роста человека при условии, что его вес равен 70 килограммам, будет равно среднему росту людей, вес которых равен 70 килограммам.

Условное математическое ожидание обозначается так:

E(Y|X=x)

где Y - случайная величина, которую мы хотим предсказать, X - случайная величина, которая нам известна, x - значение случайной величины X.

Линия регрессии - это график, который показывает зависимость между двумя случайными величинами. Она показывает, как изменяется среднее значение одной случайной величины при изменении значения другой случайной величины.

Например, рассмотрим ту же самую ситуацию с ростом и весом. Линия регрессии роста по весу покажет, как изменяется среднее значение роста человека при изменении его веса.

Уравнение линии регрессии записывается так:

y = a + bx

Y - зависимая переменная (значение, которое мы пытаемся предсказать)

X - независимая переменная (признак, который мы используем для предсказания)

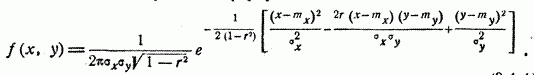
a - коэффициент наклона

b - свободный член

**34.Нормальный закон на плоскости**

Из законов распределения системы двух случайных величин имеет смысл специально рассмотреть нормальный закон, как имеющий наибольшее распространение на практике. Так как система двух случайных величин изображается случайной точкой на плоскости, нормальный закон для системы двух величин часто называют «нормальным законом на плоскости».

Закон зависит от 5 параметров: Mx My – мат ожидания, σх σу – среднеквадтичные отклонения, r – коэф. Корреляции.



Нормальный закон на плоскости нужен для описания распределения вероятностей двух случайных величин. Он имеет ряд важных свойств, которые делают его удобным для анализа и моделирования случайных явлений.

Во-первых, нормальный закон является симметричным, то есть его плотность распределения имеет форму колокола с центром в среднем значении обеих величин. Это означает, что наиболее вероятные значения находятся около среднего, а менее вероятные - дальше от него.

Во-вторых, нормальный закон является непрерывным, то есть он допускает любые значения величин. Это означает, что вероятность любого конкретного значения равна нулю, но вероятность попадания в любой интервал значений не равна нулю.

В-третьих, нормальный закон имеет свойство сходимости, то есть при достаточно большом числе измерений распределение вероятностей суммы двух случайных величин стремится к нормальному распределению. Это свойство позволяет использовать нормальный закон для анализа результатов измерений, полученных из различных источников.

**35.Задачи математической статистики. Генеральная и выборочная совокупности. Вариационный ряд. Статистический ряд. Полигон и гистограмма. Эмпирическая функция распределения и ее свойства.**

Задачи математической статистики

Математическая статистика - это раздел математики, изучающий методы сбора, обработки и анализа статистических данных. Статистические данные - это сведения о значениях случайных величин, полученных в результате наблюдений или экспериментов.

Основные задачи математической статистики:

* Описание статистических данных. Эта задача заключается в представлении статистических данных в удобной для анализа форме. Для этого используются различные методы группировки, табуляции и графической интерпретации данных.
* Оценка параметров распределения. Эта задача заключается в приближенном определении параметров (характеристик) распределения случайной величины, по данным выборки.
* Проверка статистических гипотез. Эта задача заключается в проверке предположения о том, что распределение случайной величины соответствует определенному закону распределения.

Генеральная и выборочная совокупности

Генеральная совокупность - это совокупность всех возможных значений случайной величины. Выборочная совокупность - это совокупность значений случайной величины, полученных в результате наблюдений или экспериментов.

Вариационный ряд

Вариационный ряд - это упорядоченный по возрастанию или убыванию список значений случайной величины.

Статистический ряд

Статистический ряд - это вариационный ряд, дополненный сведениями о частоте встречаемости каждого значения случайной величины.

Полигон и гистограмма

Полигон - это графическое представление статистического ряда, построенное путем соединения последовательных точек, соответствующих средним значениям интервалов.

Гистограмма - это графическое представление статистического ряда, построенное путем объединения значений случайной величины в интервалы и нанесения на оси абсцисс границ интервалов, а на оси ординат - накопленных частот.

Эмпирическая функция распределения

Эмпирическая функция распределения - это функция, определяющая для каждого значения случайной величины вероятность того, что случайное значение будет меньше или равно этому значению.

Свойства эмпирической функции распределения

* Эмпирическая функция распределения является непрерывной на интервале, содержащем все значения генеральной совокупности.
* Эмпирическая функция распределения монотонно возрастает на этом интервале.
* В точке левая и правая производные эмпирической функции распределения равны вероятности того, что случайное значение будет равно этому значению.

Примеры задач математической статистики

* Задача о описании статистических данных. Пусть имеются данные о росте 100 студентов. Требуется построить вариационный ряд и статистический ряд.
* Задача об оценке параметров распределения. Пусть имеется выборка из 100 наблюдений нормально распределенной случайной величины. Требуется оценить среднее и дисперсию этой случайной величины.
* Задача о проверке статистических гипотез. Пусть имеется выборка из 100 наблюдений нормально распределенной случайной величины. Требуется проверить гипотезу о том, что среднее этой случайной величины равно 50, при уровне значимости 0,05.

Решение задач математической статистики

Решение задач математической статистики требует использования различных методов, таких как:

* Группировка данных.
* Табуляция данных.
* Графическая интерпретация данных.
* Методы оценки параметров распределения.
* Методы проверки статистических гипотез.

Для решения задач математической статистики используются различные компьютерные программы, такие как Microsoft Excel, SPSS, SAS и другие.

**36.Точечное оценивание параметров распределения. Свойства точечных оценок. Несмещенные оценки математического ожидания и дисперсии.**

Точечное оценивание параметров распределения - это процедура нахождения приближенного значения неизвестного параметра распределения случайной величины по данным выборки.

Свойства точечных оценок:

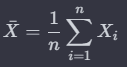
Несмещенность - математическое ожидание точечной оценки равно оцениваемому параметру.

Эффективность - точечная оценка обладает минимальной дисперсией среди всех несмещенных оценок данного параметра.

Согласованность - при увеличении объема выборки точечная оценка сходится к оцениваемому параметру по вероятности.

**Несмещенная оценка математического ожидания (среднего):**

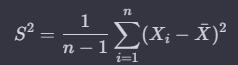
Если у нас есть выборка Х1, Х2 … Хn из случайной величины X, то несмещенная оценка математического ожидания обозначается как x̄ и вычисляется по формуле:



Эта оценка несмещенная, потому что ее математическое ожидание равно истинному математическому ожиданию случайной величины.

**Несмещенная оценка дисперсии:**

Для несмещенной оценки дисперсии используется формула:



Здесь S^2 - несмещенная оценка дисперсии. Важно отметить, что в знаменателе используется n−1 вместо n. Это делается для коррекции смещения в оценке, и она становится несмещенной.

Такие несмещенные оценки важны, чтобы приблизить выборочные статистики к параметрам генеральной совокупности, особенно при работе с ограниченными выборками данных.

**37. Интервальные оценки параметров генеральной совокупности. Доверительная вероятность.**

Интервальная оценка - это оценка параметра генеральной совокупности, которая представляет собой интервал, в котором с заданной вероятностью находится истинное значение параметра.

Доверительная вероятность - это вероятность того, что истинное значение параметра генеральной совокупности попадает в доверительный интервал.

Например, если доверительный интервал для среднего возраста составляет от 20 до 25 лет с доверительной вероятностью 0,95, то это означает, что с вероятностью 0,95 истинное среднее значение возраста жителей данного города находится в пределах от 20 до 25 лет.

Виды интервальных оценок

Существует множество способов построения интервальных оценок. Наиболее распространенными являются:

Интервальные оценки для среднего значения

Для построения интервальной оценки среднего значения используется формула:

x̄ ± t⋅s/√n

где x̄ - выборочное среднее, t - значение критического значения t-критерия Стьюдента, зависящее от доверительной вероятности и объема выборки, s - выборочное стандартное отклонение, n - выборка.

Например, если доверительный интервал для среднего значения составляет от 20 до 25 лет с доверительной вероятностью 0,95, а объем выборки составляет 100 человек, то значение критического значения t-критерия Стьюдента составляет 1,96. В этом случае доверительный интервал будет иметь следующий вид:

x̄ ± 1,96⋅s/√100

Доверительная вероятность - это вероятность того, что истинное значение параметра генеральной совокупности попадает в доверительный интервал.

Доверительный интервал - это интервал, в котором с заданной вероятностью находится истинное значение параметра.

На практике обычно используются следующие значения доверительной вероятности:

0,95 (95%-ный доверительный интервал)

0,99 (99%-ный доверительный интервал)

0,999 (99,9%-ный доверительный интервал)

Преимущества и недостатки интервальных оценок

Преимущества интервальных оценок:

Они дают более полную информацию о неизвестном параметре, чем точечные оценки.

Они позволяют оценить надежность оценки.

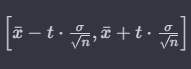
Недостатки интервальных оценок:

Они более сложны в интерпретации, чем точечные оценки.

Они менее точные, чем точечные оценки, построенные на больших выборках.

**38. Построение доверительного интервала для математического ожидания нормально распределенной генеральной совокупности.**

Доверительный интервал для математического ожидания нормально распределенной генеральной совокупности строится по следующей формуле:



Где x̄ – выборчное среднее

t – значение t-статистики, определённое уровнем доверия

σ - стандартное отклонение генеральной совокупности

n - размер выборки

Например, предположим, что мы имеем выборку из 100 наблюдений нормально распределенной генеральной совокупности с известным средним квадратическим отклонением σ = 10. Мы хотим построить доверительный интервал для математического ожидания генеральной совокупности с доверительной вероятностью 0,95.

Критическое значение распределения Стьюдента при доверительной вероятности 0,95 и степени свободы n - 1 = 95 составляет 1,96.

Таким образом, доверительный интервал для математического ожидания генеральной совокупности будет иметь следующий вид:

[x̄ - 1,96\*10/√100, x̄ + 1,96\*10/√100]

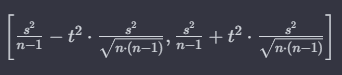
Подставив в эту формулу значение выборочного среднего x̄ = 105, получим

[104,04, 105,96]

Следовательно, с доверительной вероятностью 0,95 мы можем утверждать, что математическое ожидание генеральной совокупности находится в интервале от 104,04 до 105,96.

**39. Построение доверительного интервала для дисперсии нормально распределенной генеральной совокупности.**

Доверительный интервал для дисперсии нормально распределенной генеральной совокупности строится по следующей формуле:



Где: s^2 - выборочная дисперсия,

n - размер выборки,

t - значение t-статистики, определенное уровнем доверия.

Например, предположим, что мы имеем выборку из 100 наблюдений нормально распределенной генеральной совокупности с выборочной дисперсией s^2 = 120. Мы хотим построить доверительный интервал для дисперсии генеральной совокупности с доверительной вероятностью 0,95.

Критическое значение распределения Стьюдента при доверительной вероятности 0,95 и степени свободы n - 1 = 95 составляет 1,96.

Подставив в эту формулу значение выборочной дисперсии s^2 = 120, получим:

[12,12, 127,88]

**40. Основные понятия теории проверки гипотез. Простая и сложная гипотезы. Нулевая и альтернативная гипотезы. Статистический критерий. Область принятия гипотезы и критическая область. Ошибки первого и второго родов. Уровень значимости и мощность критерия. Двусторонняя и односторонняя критические области.**

Теория проверки гипотез — это важный компонент статистики, используемый для сделки выводов о популяции на основе выборочных данных. Вот основные понятия:

Основные понятия теории проверки гипотез

Теория проверки гипотез - это раздел математической статистики, который занимается разработкой методов для принятия решений о справедливости статистических гипотез.

*Простая и сложная гипотезы*

Простая гипотеза - это гипотеза, которая может быть сформулирована как равенство двух числовых значений, например, H0: μ = 100.

Сложная гипотеза - это гипотеза, которая не является простой, например, H0: μ > 100.

*Нулевая и альтернативная гипотезы*

Нулевая гипотеза (H0) - это гипотеза, которую мы пытаемся опровергнуть, например, H0: μ = 100.

Альтернативная гипотеза (H1) - это гипотеза, которую мы пытаемся принять, если нулевая гипотеза опровергнута, например, H1: μ > 100.

*Статистический критерий*

Статистический критерий - это функция выборки, которая используется для принятия решения о справедливости гипотезы.

*Область принятия гипотезы и критическая область*

Область принятия гипотезы - это область значений критерия, при которых нулевая гипотеза не отвергается.

Критическая область - это область значений критерия, при которых нулевая гипотеза отвергается.

*Ошибки первого и второго родов*

Ошибка первого рода - это ситуация, когда нулевая гипотеза отвергается, хотя она верна.

Ошибка второго рода - это ситуация, когда нулевая гипотеза не отвергается, хотя она неверна.

*Уровень значимости и мощность критерия*

Уровень значимости - это вероятность совершить ошибку первого рода.

Мощность критерия - это вероятность отвергнуть нулевую гипотезу, когда она неверна.

*Двусторонняя и односторонняя критические области*

Двусторонняя критическая область - это область значений критерия, при которых нулевая гипотеза отвергается, независимо от того, больше или меньше выборочное значение критического значения.

Односторонняя критическая область - это область значений критерия, при которых нулевая гипотеза отвергается только в том случае, если выборочное значение больше или меньше критического значения в определенном направлении.

**41. Проверка гипотезы о виде закона распределения. Критерий согласия χ2 Пирсона.**

Критерий согласия Пирсона - это непараметрический критерий, который используется для проверки гипотезы о виде закона распределения генеральной совокупности.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: генеральная совокупность распределена по закону распределения F(x).

Альтернативная гипотеза (H1) может быть сформулирована следующим образом:

H1: генеральная совокупность не распределена по закону распределения F(x).

Статистический критерий критерия согласия Пирсона вычисляется по следующей формуле:

χ^2 = ∑(Oi - Ei)^2 / Ei

где:

Oi - наблюдаемая частота в i-й категории

Ei - ожидаемая частота в i-й категории, вычисляемая по закону распределения F(x)

Область принятия гипотезы критерия согласия Пирсона определяется критическим значением χ^2, которое зависит от объема выборки n, числа категорий k и уровня значимости α.

Если значение критерия χ^2 попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

Ошибки первого и второго рода

В случае критерия согласия Пирсона ошибка первого рода - это ситуация, когда нулевая гипотеза отвергается, хотя она верна. Вероятность совершить ошибку первого рода определяется уровнем значимости α.

Ошибка второго рода - это ситуация, когда нулевая гипотеза не отвергается, хотя она неверна. Вероятность совершить ошибку второго рода зависит от распределения генеральной совокупности.

Мощность критерия

Мощность критерия согласия Пирсона - это вероятность отвергнуть нулевую гипотезу, когда она неверна. Мощность критерия зависит от распределения генеральной совокупности, уровня значимости α и объема выборки n.

**42. Критерии значимости. Проверка гипотез о математических ожиданиях одной и двух независимых нормальных выборок.**

Проверка гипотез о математических ожиданиях одной нормальной выборки

Для проверки гипотезы о математическом ожидании одной нормальной выборки используется t-критерий Стьюдента.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: μ = μ0

Альтернативная гипотеза (H1) может быть сформулирована следующим образом:

H1: μ ≠ μ0

Статистический критерий t-критерия Стьюдента вычисляется по следующей формуле:

t = (x̄ - μ0) / s / √n

где:

x̄ - выборочное среднее

μ0 - значение математического ожидания, которое проверяется

s - выборочная дисперсия

n - объем выборки

Область принятия гипотезы t-критерия Стьюдента определяется критическим значением t, которое зависит от объема выборки n, уровня значимости α и дисперсии генеральной совокупности σ^2.

Если значение критерия t попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

**43. Критерии значимости. Проверка гипотез о дисперсиях одной и двух независимых нормальных выборок.**

Проверка гипотез о дисперсиях одной нормальной выборки

Для проверки гипотезы о дисперсии одной нормальной выборки используется F-критерий Фишера.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: σ^2 = σ^20

Альтернативная гипотеза (H1) может быть сформулирована следующим образом:

H1: σ^2 ≠ σ^20

Статистический критерий F-критерия Фишера вычисляется по следующей формуле:

F = s^2 / σ^20

где:

s^2 - выборочная дисперсия

σ^20 - значение дисперсии, которое проверяется

Область принятия гипотезы F-критерия Фишера определяется критическим значением F, которое зависит от объема выборки n и уровня значимости α.

Если значение критерия F попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

**44. Критерии значимости. Проверка гипотез о математических ожиданиях двух зависимых и независимых нормальных выборок.**

Проверка гипотез о математических ожиданиях двух зависимых нормальных выборок

Для проверки гипотезы о равенстве математических ожиданий двух зависимых нормальных выборок используется t-критерий Стьюдента для зависимых выборок.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: μ1 = μ2

Альтернативная гипотеза (H1) может быть сформулирована следующим образом:

H1: μ1 ≠ μ2

Статистический критерий t-критерия Стьюдента для зависимых выборок вычисляется по следующей формуле:

t = (x̄1 - x̄2) / s / √n

где:

x̄1 - выборочное среднее первой выборки

x̄2 - выборочное среднее второй выборки

s - выборочное стандартное отклонение, которое рассчитывается по формуле:

s = √(s^21 + s^22) / 2

где:

s^21 - выборочная дисперсия первой выборки

s^22 - выборочная дисперсия второй выборки

Область принятия гипотезы t-критерия Стьюдента для зависимых выборок определяется критическим значением t, которое зависит от объема выборки n и уровня значимости α.

Если значение критерия t попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

Пример использования t-критерия Стьюдента для зависимых выборок

Предположим, что мы имеем две зависимые выборки из 10 наблюдений, нормально распределенных. Нулевая гипотеза H0 гласит, что математические ожидания двух выборок равны.

Для проверки этой гипотезы вычислим t-критерий Стьюдента для зависимых выборок:

t = (x̄1 - x̄2) / s / √n = (101 - 99) / √(10^2 + 10^2) / 10 = 1,414

Критическое значение t для уровня значимости α = 0,05 и объема выборки n = 10 равно 2,262.

Так как значение критерия t (1,414) не попадает в критическую область, то нулевая гипотеза не отвергается. С вероятностью 95% мы можем утверждать, что математические ожидания двух выборок равны.

**45. Использование распределения Стьюдента при построении доверительных интервалов и проверке статистических гипотез.**

Распределение Стьюдента - это распределение вероятностей, которое используется для построения доверительных интервалов и проверки статистических гипотез в том случае, когда распределение генеральной совокупности неизвестно или предполагается нормальным.

Построение доверительных интервалов

Доверительный интервал - это интервал, в котором с заданной вероятностью находится истинное значение параметра генеральной совокупности.

Для построения доверительного интервала для математического ожидания генеральной совокупности, распределенной нормально, используется следующая формула:

x̄ ± tα/2 \* s / √n

где:

x̄ - выборочное среднее

tα/2 - критическое значение распределения Стьюдента с уровнем значимости α/2

s - выборочное стандартное отклонение

n - объем выборки

Например, если мы хотим построить доверительный интервал для математического ожидания генеральной совокупности с уровнем значимости α = 0,05 и объемом выборки n = 100, то критическое значение распределения Стьюдента tα/2 = 1,96.

Если выборочное среднее равно x̄ = 100, а выборочное стандартное отклонение равно s = 10, то доверительный интервал будет следующим:

100 ± 1,96 \* 10 / √100 = 100 ± 1,96

Это означает, что с вероятностью 95% истинное значение математического ожидания генеральной совокупности находится в интервале от 98,04 до 101,96.

Проверка статистических гипотез

Для проверки статистических гипотез о математическом ожидании генеральной совокупности используется t-критерий Стьюдента.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: μ = μ0

Альтернативная гипотеза (H1) может быть сформулирована следующим образом:

H1: μ ≠ μ0

Статистический критерий t-критерия Стьюдента вычисляется по следующей формуле:

t = (x̄ - μ0) / (s / √n)

где:

x̄ - выборочное среднее

μ0 - значение математического ожидания, которое проверяется

s - выборочное стандартное отклонение

n - объем выборки

Если значение критерия t попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

**46. Использование нормального распределения при построении доверительных интервалов и проверке статистических гипотез.**

Нормальное распределение - это непрерывное распределение вероятностей, которое имеет колоколообразную форму.

Построение доверительных интервалов

Доверительный интервал - это интервал, в котором с заданной вероятностью находится истинное значение параметра генеральной совокупности.

Для построения доверительного интервала для математического ожидания генеральной совокупности, распределенной нормально, используется следующая формула:

x̄ ± zα/2 \* σ / √n

где:

x̄ - выборочное среднее

zα/2 - критическое значение нормального распределения с уровнем значимости α/2

σ - известное значение генеральной дисперсии

n - объем выборки

Например, если мы хотим построить доверительный интервал для математического ожидания генеральной совокупности с уровнем значимости α = 0,05 и объемом выборки n = 100, а известное значение генеральной дисперсии равно σ = 10, то критическое значение нормального распределения zα/2 = 1,96.

Если выборочное среднее равно x̄ = 100, то доверительный интервал будет следующим:

100 ± 1,96 \* 10 / √100 = 100 ± 1,96

Это означает, что с вероятностью 95% истинное значение математического ожидания генеральной совокупности находится в интервале от 98,04 до 101,96.

Проверка статистических гипотез

Для проверки статистических гипотез о математическом ожидании генеральной совокупности используется t-критерий Стьюдента.

Нулевая гипотеза (H0) в этом случае формулируется следующим образом:

H0: μ = μ0

Альтернативная гипотеза может быть сформулирована следующим образом:

H1: μ ≠ μ0

Статистический критерий t-критерия Стьюдента вычисляется по следующей формуле:

t = (x̄ - μ0) / (σ / √n)

где:

x̄ - выборочное среднее

μ0 - значение математического ожидания, которое проверяется

σ - известное значение генеральной дисперсии

n - объем выборки

Если значение критерия t попадает в критическую область, то нулевая гипотеза отвергается. В противном случае нулевая гипотеза не отвергается.

**47. Использование χ2 -распределения при построении доверительных интервалов и проверке статистических гипотез.**

Двуквадратное распределение - это распределение вероятностей, которое используется для построения доверительных интервалов и проверки статистических гипотез в том случае, когда распределение генеральной совокупности неизвестно или предполагается нормальным.

Построение доверительных интервалов

Доверительный интервал - это интервал, в котором с заданной вероятностью находится истинное значение параметра генеральной совокупности.

Для построения доверительного интервала для дисперсии генеральной совокупности, распределенной нормально, используется следующая формула:

σ^2 = (x̄ - μ)^2 / χ^2α/2,n-1

где:

x̄ - выборочное среднее

μ - неизвестное значение математического ожидания генеральной совокупности

χ^2α/2,n-1 - критическое значение двуквадратного распределения с уровнем значимости α/2 и числом степеней свободы n-1

Например, если мы хотим построить доверительный интервал для дисперсии генеральной совокупности с уровнем значимости α = 0,05 и объемом выборки n = 100, то критическое значение двуквадратного распределения χ^2α/2,n-1 = 38,384.

Если выборочное среднее равно x̄ = 100, а выборочная дисперсия равна s^2 = 10, то доверительный интервал будет следующим:

10 = (100 - μ)^2 / 38,384,100-1 = 3.82

Это означает, что с вероятностью 95% истинная дисперсия генеральной совокупности находится в интервале от 2.18 до 5.42.

**48. Виды зависимостей между случайными величинами. Основные задачи корреляционного и регрессионного анализа.**

Виды зависимостей между случайными величинами

Зависимость между случайными величинами - это такое их свойство, при котором изменение одной величины приводит к изменению другой.

Зависимости между случайными величинами могут быть функциональными и статистическими.

Функциональная зависимость - это такая зависимость, при которой каждому значению одной величины соответствует одно и только одно значение другой величины.

Статистическая зависимость - это такая зависимость, при которой изменение одной величины приводит к изменению другой с некоторой вероятностью.

Статистическая зависимость может быть линейной и нелинейной.

Линейная зависимость - это такая зависимость, при которой зависимость между случайными величинами может быть представлена линейной функцией.

Нелинейная зависимость - это такая зависимость, при которой зависимость между случайными величинами не может быть представлена линейной функцией.

Основные задачи корреляционного и регрессионного анализа

Корреляционный анализ - это метод, позволяющий оценить силу и направление связи между двумя или более случайными величинами.

Основные задачи корреляционного анализа:

* Оценка силы и направления связи между двумя или более случайными величинами.
* Проверка гипотез о наличии связи между случайными величинами.
* Определение типа связи между случайными величинами.

Регрессионный анализ - это метод, позволяющий построить модель зависимости между одной случайной величиной (зависимой) и одной или несколькими случайными величинами (независимыми).

Основные задачи регрессионного анализа:

* Построение модели зависимости между одной случайной величиной (зависимой) и одной или несколькими случайными величинами (независимыми).
* Оценка параметров модели зависимости.
* Проверка гипотез о параметрах модели зависимости.
* Использование модели зависимости для прогнозирования значений зависимой величины.

Связь между корреляционным и регрессионным анализом

Корреляционный анализ является предварительным этапом регрессионного анализа. Корреляционный анализ позволяет оценить силу и направление связи между случайными величинами. Если связь между случайными величинами слабая или отсутствует, то регрессионный анализ нецелесообразен.

Если связь между случайными величинами достаточно сильна, то регрессионный анализ позволяет построить модель зависимости между этими случайными величинами. Модель зависимости может быть использована для прогнозирования значений зависимой величины.

**49. Выборочный коэффициент корреляции и его свойства.**

Выборочный коэффициент корреляции - это статистическая оценка коэффициента корреляции двух случайных величин.

Формула выборочного коэффициента корреляции

Выборочный коэффициент корреляции между двумя случайными величинами X и Y определяется по следующей формуле:



где:

r - выборочный коэффициент корреляции

x\_i - i-ое значение случайной величины X

y\_i - i-ое значение случайной величины Y

x̄ - выборочное среднее случайной величины X

ȳ - выборочное среднее случайной величины Y

Свойства выборочного коэффициента корреляции

*Свойство 1.* Выборочный коэффициент корреляции всегда лежит в интервале от -1 до 1.

*Свойство 2*. Если r = 1, то между случайными величинами X и Y существует полная положительная линейная зависимость.

*Свойство 3.* Если r = -1, то между случайными величинами X и Y существует полная отрицательная линейная зависимость.

*Свойство 4*. Если r = 0, то между случайными величинами X и Y отсутствует линейная зависимость.

*Свойство 5.* Чем ближе значение r к 1 или -1, тем сильнее связь между случайными величинами X и Y.

*Свойство 6.* Чем больше объем выборки, тем точнее оценка коэффициента корреляции.

Выборочный коэффициент корреляции используется для оценки силы и направления связи между двумя случайными величинами.

Если значение выборочного коэффициента корреляции близко к 1 или -1, то можно утверждать, что между случайными величинами существует сильная связь.

Если значение выборочного коэффициента корреляции близко к 0, то можно утверждать, что между случайными величинами отсутствует связь.

Однако следует отметить, что выборочный коэффициент корреляции является статистической оценкой, и его значение может случайно отклоняться от истинного значения коэффициента корреляции генеральной совокупности.

Поэтому для оценки значимости связи между случайными величинами необходимо использовать критерий корреляции Пирсона.

**50. Эмпирическое линейное уравнение регрессии. Метод наименьших квадратов.**

Эмпирическое линейное уравнение регрессии строится с использованием метода наименьших квадратов (МНК), который минимизирует сумму квадратов разностей между фактическими значениями зависимой переменной и предсказанными значениями линейной модели.

Предположим, что у нас есть пары значений (X, Y), где X - независимая переменная, Y - зависимая переменная. Мы хотим найти линейное уравнение вида Y=a+bX, которое наилучшим образом подходит к данным.

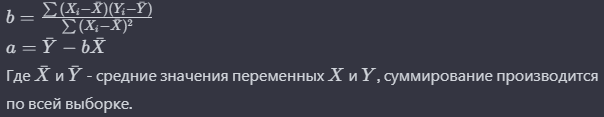
Этапы построения эмпирического линейного уравнения регрессии методом наименьших квадратов:

1) Определение уравнения:

Форма уравнения: Y=a+bX, где a - точка пересечения с осью Y, b - коэффициент наклона.

2) Расчет коэффициентов a и b:

Коэффициенты a и b вычисляются по следующим формулам:



3) Построение уравнения регрессии:

Подставим найденные значения a и b в уравнение, получив тем самым эмпирическое линейное уравнение регрессии.

4) Интерпретация коэффициентов:

Коэффициент a представляет собой значение Y, когда X равен 0.

Коэффициент b представляет собой изменение в Y за единичное изменение в X.

5) Проверка адекватности модели:

После построения модели важно оценить ее адекватность с использованием статистических тестов и графиков

Метод наименьших квадратов обеспечивает оптимальные коэффициенты, минимизируя сумму квадратов остатков (разностей между фактическими и предсказанными значениями). Это обеспечивает наилучшее приближение линейной модели к данным.