# 

[**ЛЕКЦИЯ 1**](#_w5ya9pj8la61) **1**

[**Лекция 2. Выборка. Эмпирическая функция. Выборочные моменты**](#_d1c4e1umgmgq) **18**

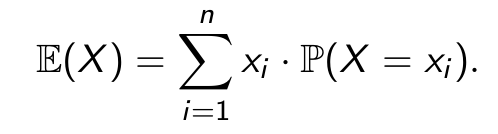
# 

# **ЛЕКЦИЯ 1**

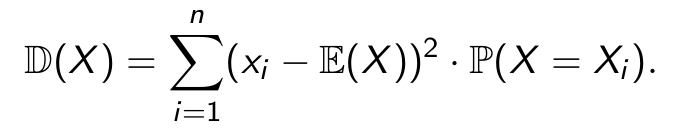
**ÔN TẬP VỀ LÝ THUYẾT XÁC SUẤT - CÁC VẤN ĐỀ VỀ THỐNG KÊ TOÁN HỌC**

1. Đặc tính số của các biến ngẫu nhiên rời rạc.

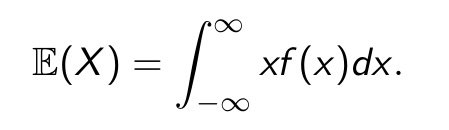
Kỳ vọng toán học:



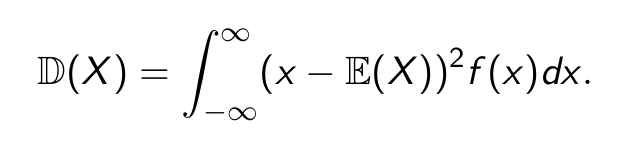
Phương sai:



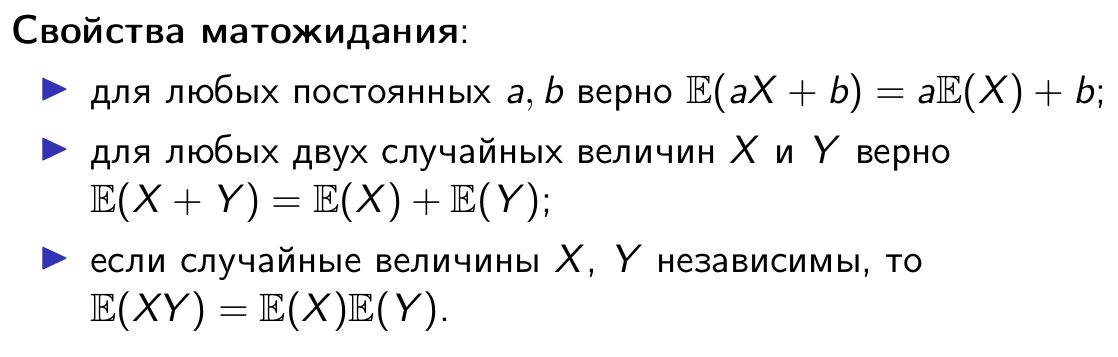
1. Đặc tính số của biến ngẫu nhiên liên tục

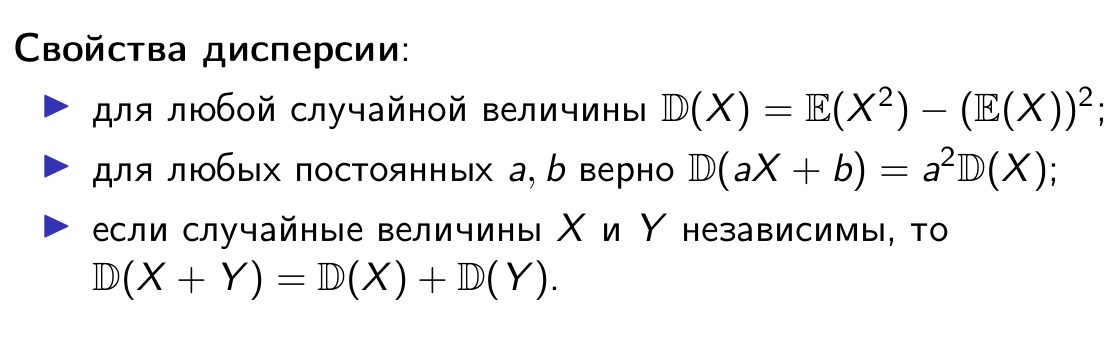
Kỳ vọng toán học:

Phương sai:



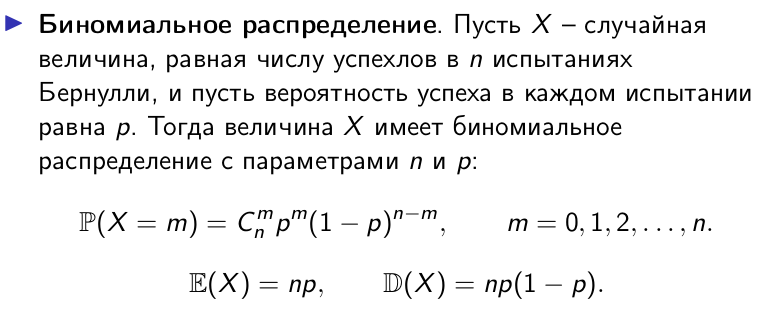
1. Đặc tính số của một biến ngẫu nhiên





1. **Phân phối cơ bản**

* **Phân phối nhị thức:** Gọi X là biến ngẫu nhiên bằng số lần thành công trong n lần thử Bernoulli và đặt xác suất thành công trong mỗi lần thử là p. Khi đó đại lượng X có phân bố nhị thức với tham số n và p:

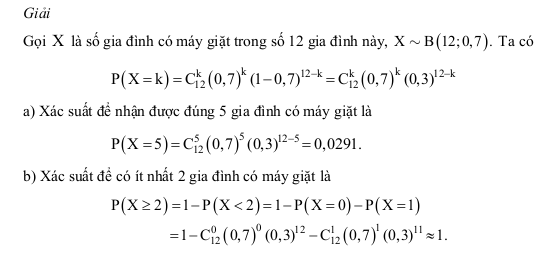


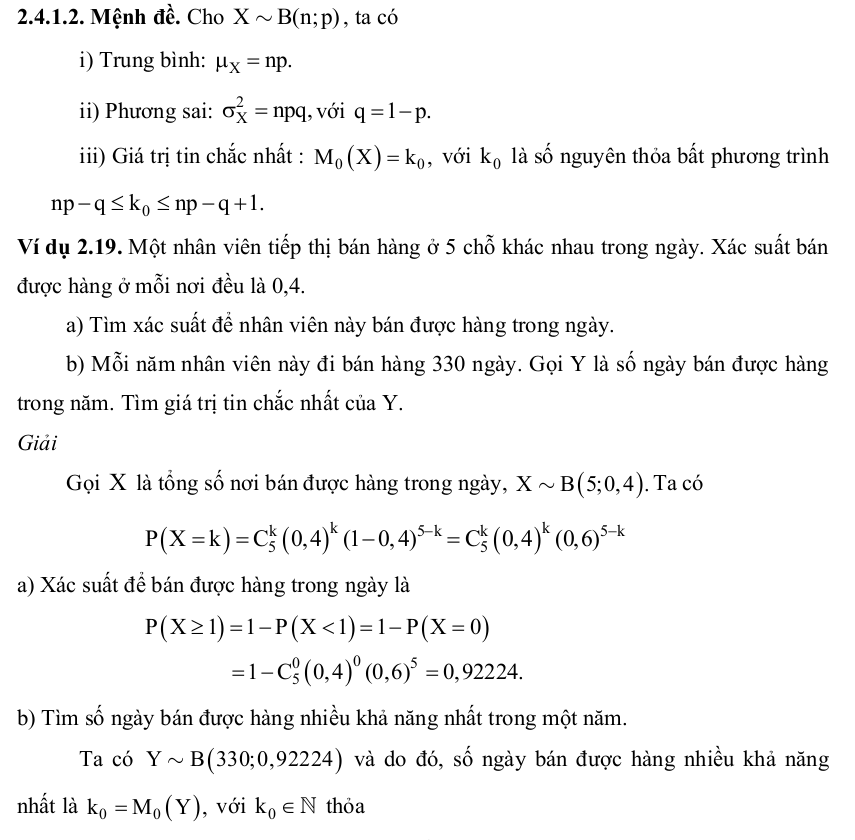
Ví dụ 2.18. Trong một vùng dân cư có 70% gia đình có máy giặt, chọn ngẫu nhiên 12 gia

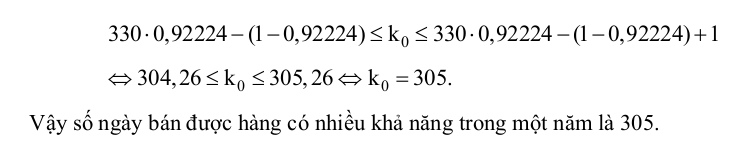
đình. Tính xác suất

a) có đúng 5 gia đình có máy giặt.

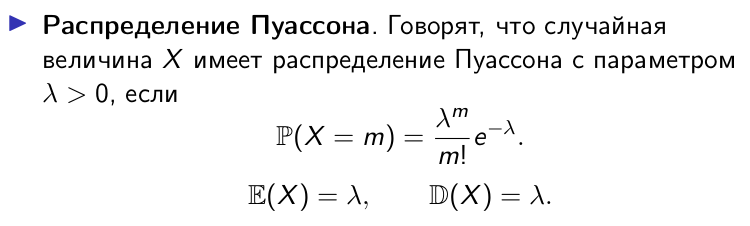
b) có ít nhất 2 gia đình có máy giặt.

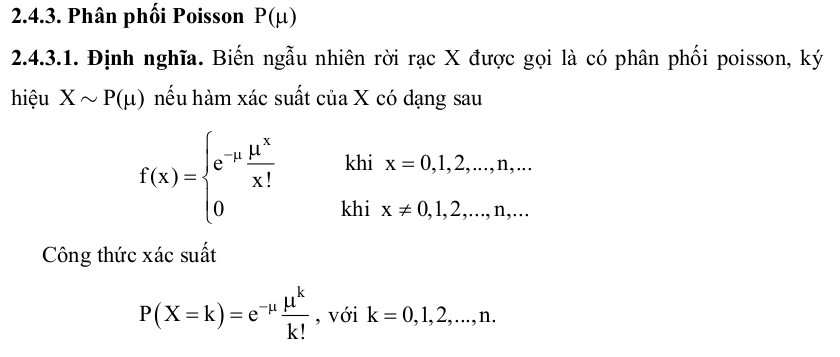


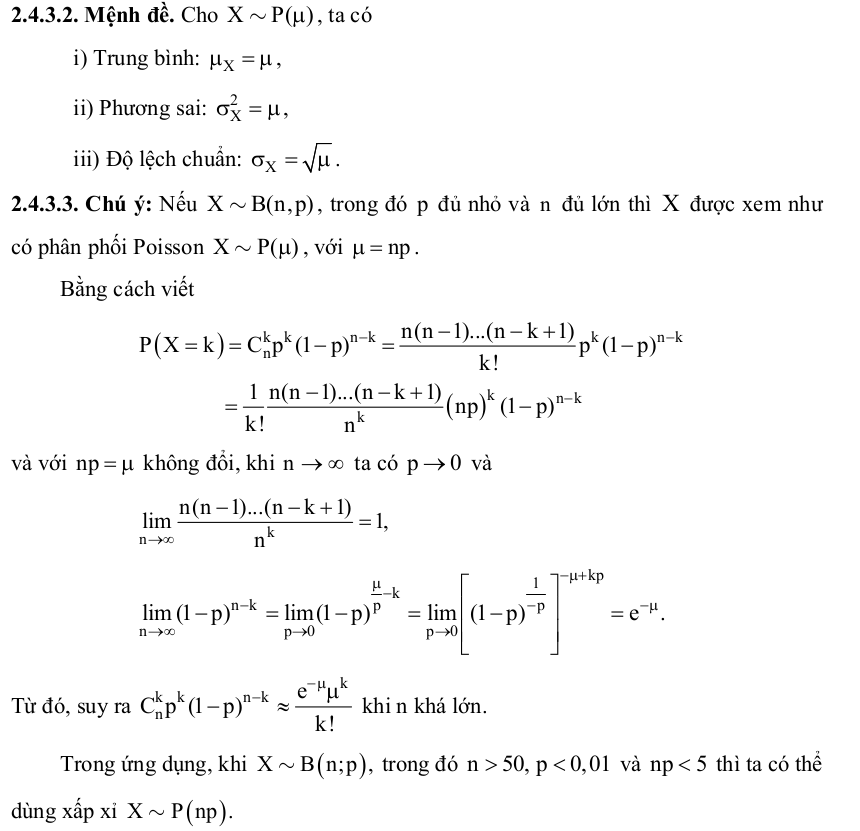


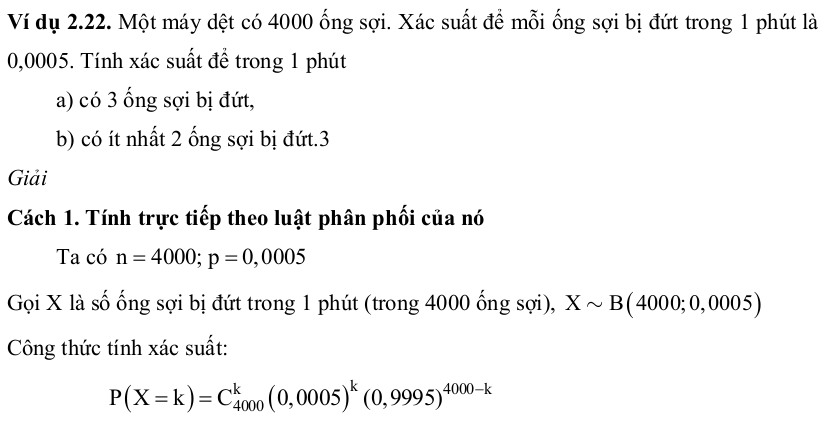


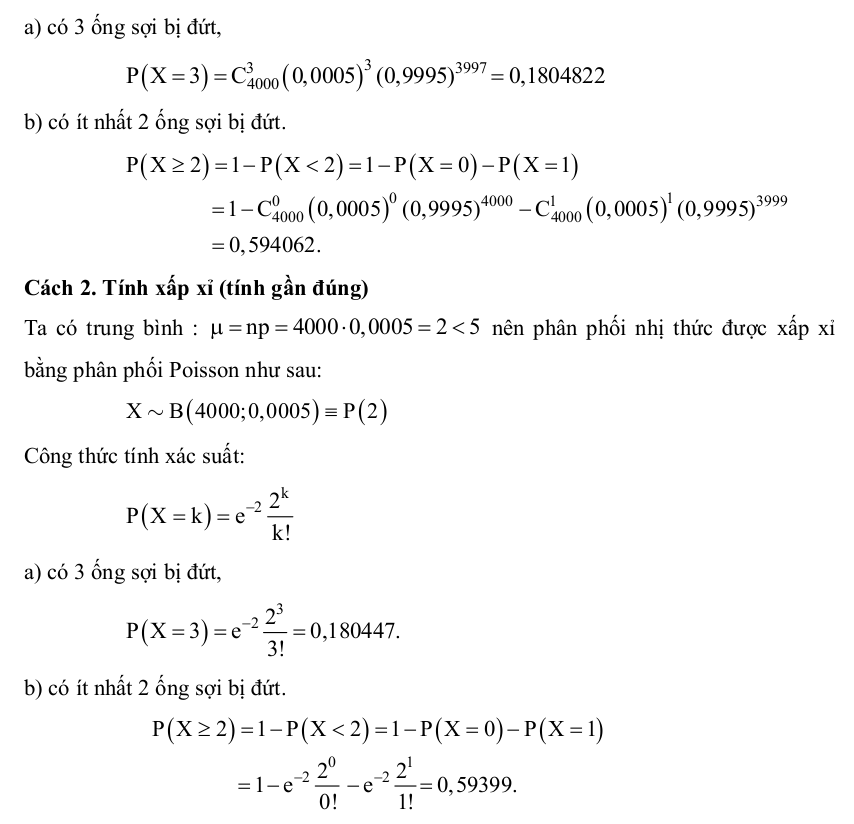
* **Phân phối Poisson: Một biến ngẫu nhiên X được gọi là có phân phối Poisson với tham số λ > 0 nếu**

****

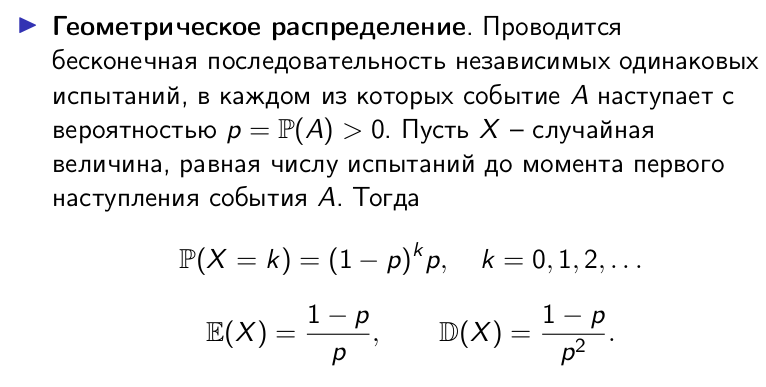
****

****

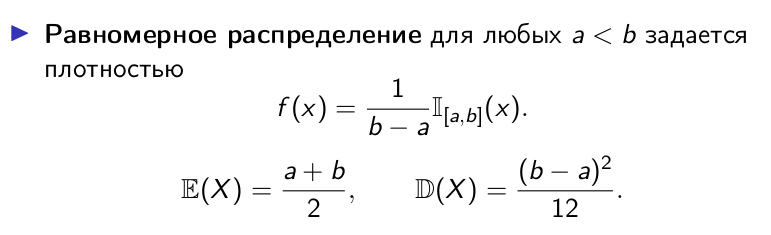
****

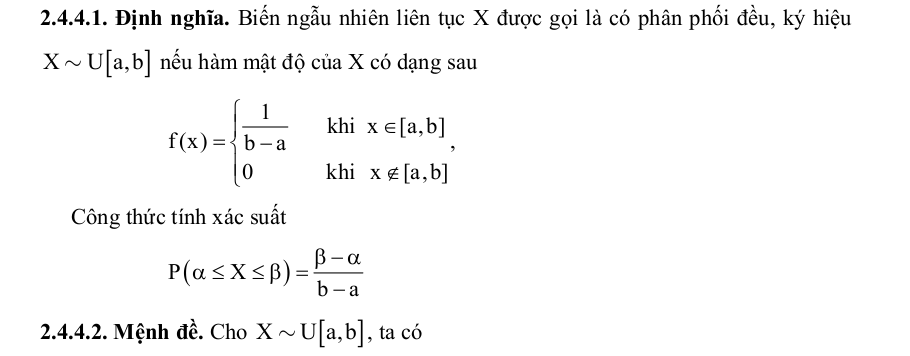
****

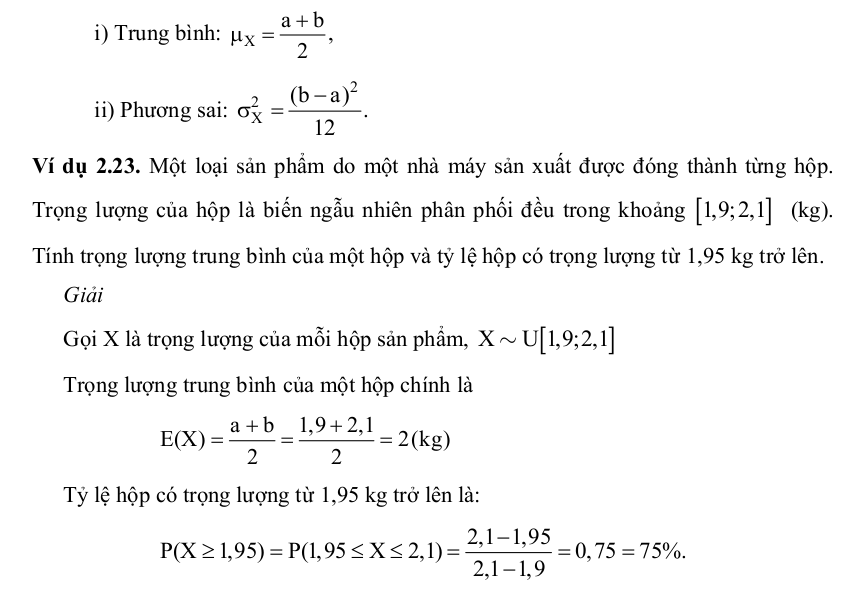
* **Phân phối hình học: Một chuỗi vô hạn các phép thử độc lập giống hệt nhau được thực hiện, trong đó mỗi sự kiện A xảy ra với xác suất p = P(A) > 0. Giả sử X là một biến ngẫu nhiên bằng số lần thử trước khi sự kiện A xảy ra lần đầu tiên. Khi đó**

****

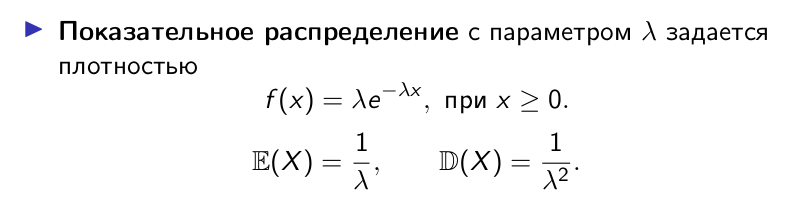
* **Một phân phối đồng đều cho mọi a < b được cho bởi mật độ**

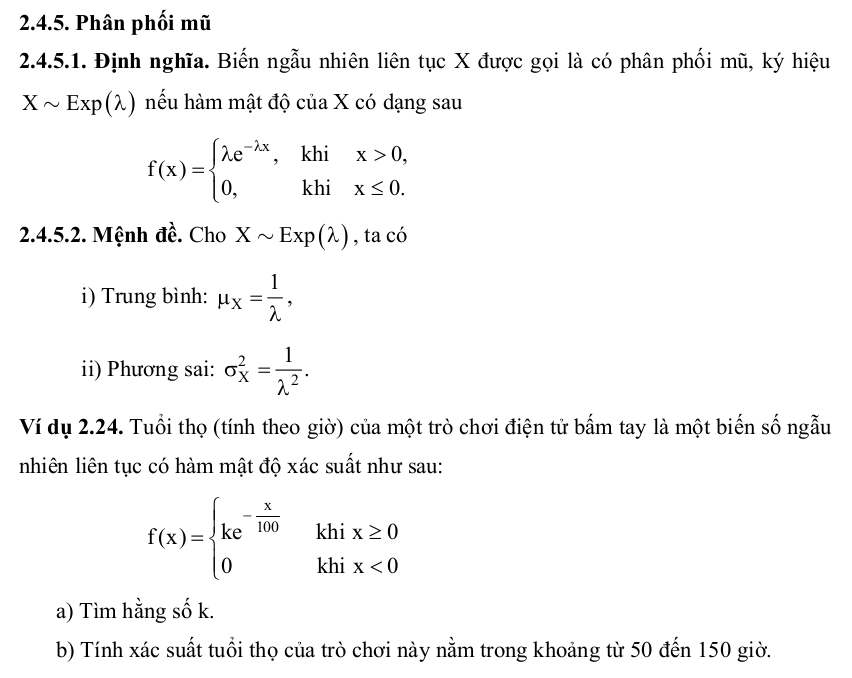


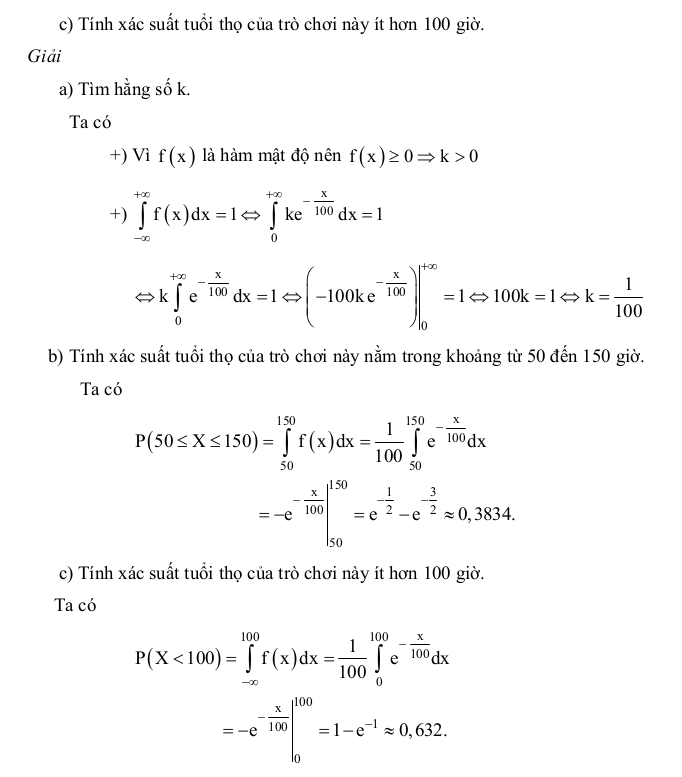




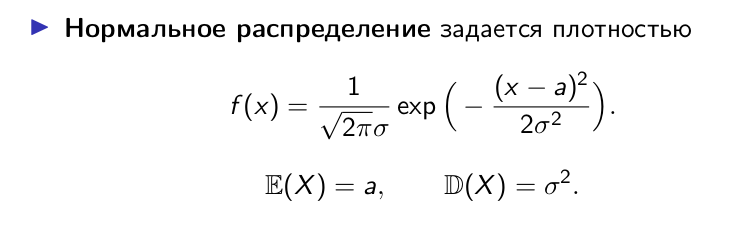
* **Phân phối hàm mũ với tham số λ được cho bởi mật độ**

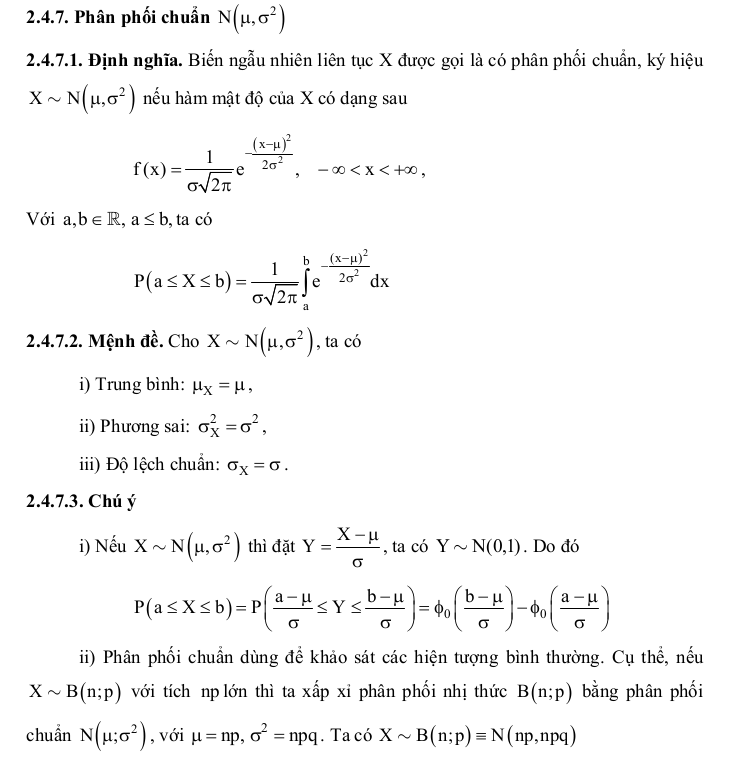
****

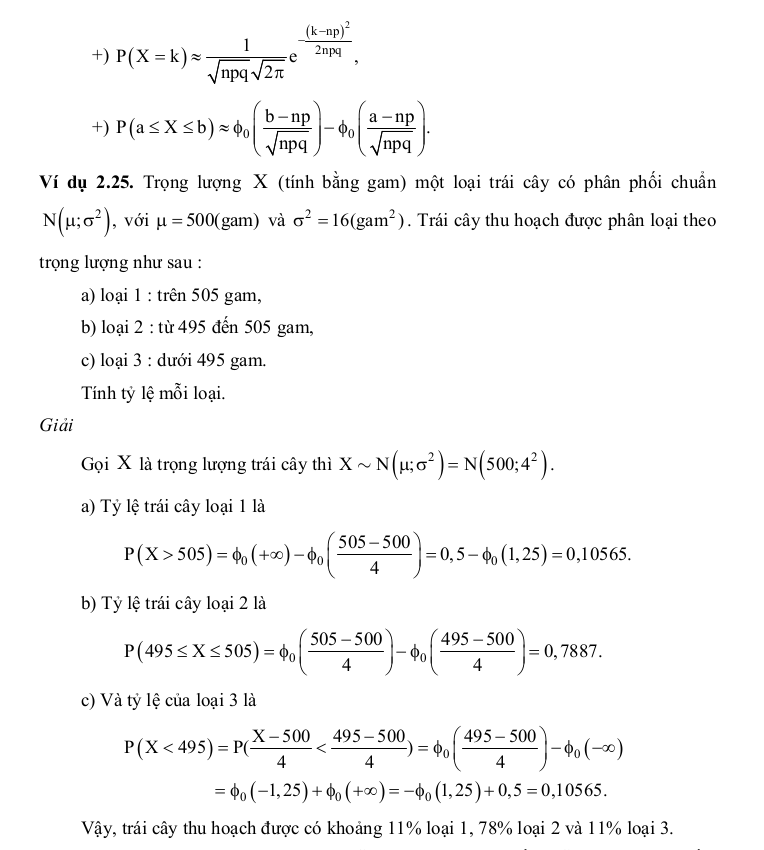
****

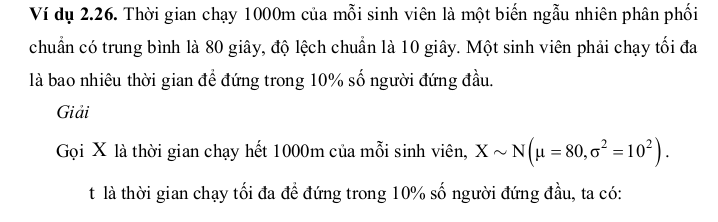
****

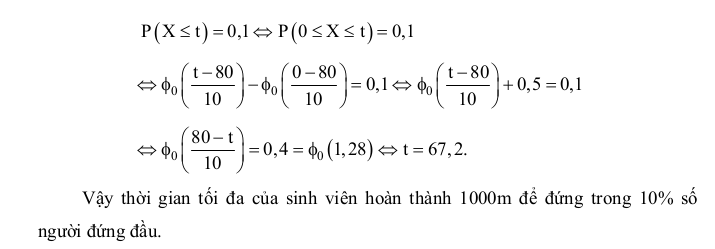
* **Sự phân phối chuẩn được cho bởi mật độ**

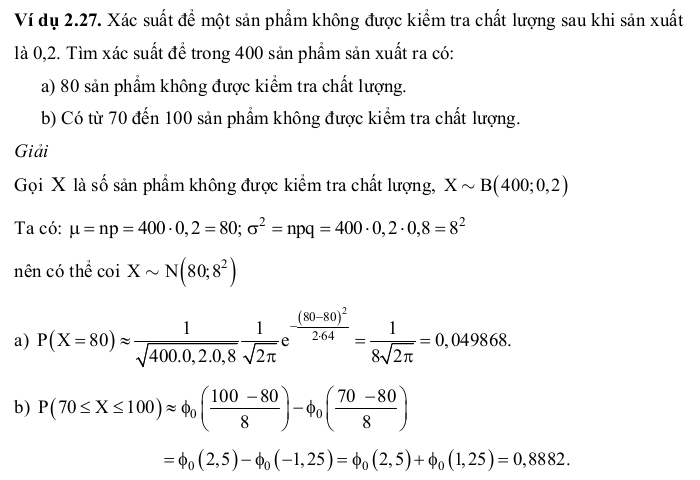
****





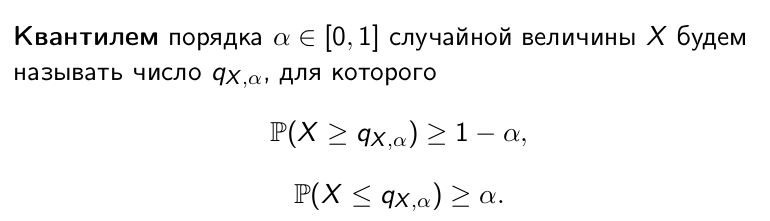






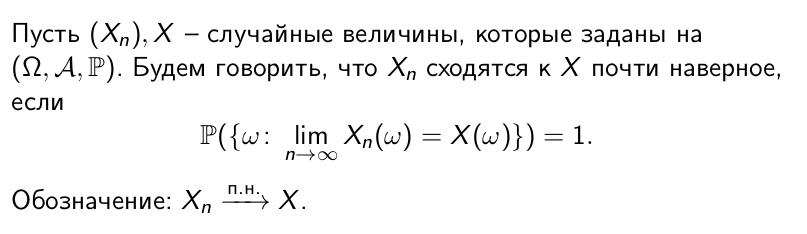
1. **Квантиль - Lượng tử**

**Một lượng tử bậc α ∈ [0, 1] của biến ngẫu nhiên X là số qX ,α mà**

****

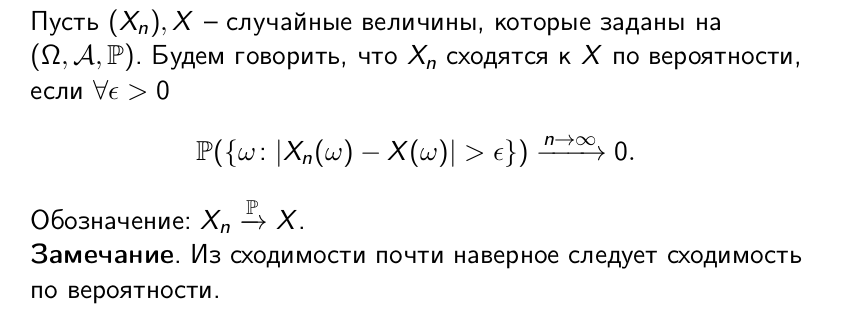
1. **Вероятностные сходимости - sự hội tụ xác suất**

**Đặt (Xn ), X là các biến ngẫu nhiên được xác định trên (Ω, A, P). Ta nói rằng Xn gần như chắc chắn hội tụ về X nếu**

****

**Đặt (Xn ), X là các biến ngẫu nhiên được xác định trên (Ω, A, P). Chúng ta sẽ nói rằng Xn hội tụ về X có xác suất nếu ∀ϵ > 0**

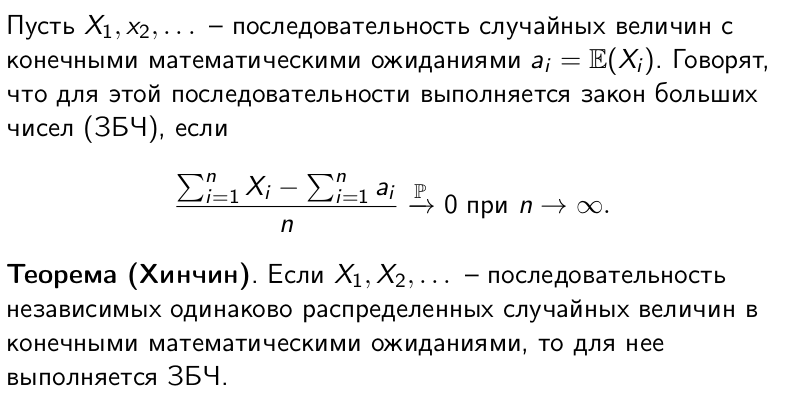
**Пусть (Xn ), X – случайные величины, которые заданы на (Ω, A, P). Будем говорить, что Xn сходятся к X по вероятности, если ∀ϵ > 0**

****

**Bình luận. Sự hội tụ gần như chắc chắn hàm ý sự hội tụ về xác suất**

1. **Закон больших чисел - Luật số lớn**

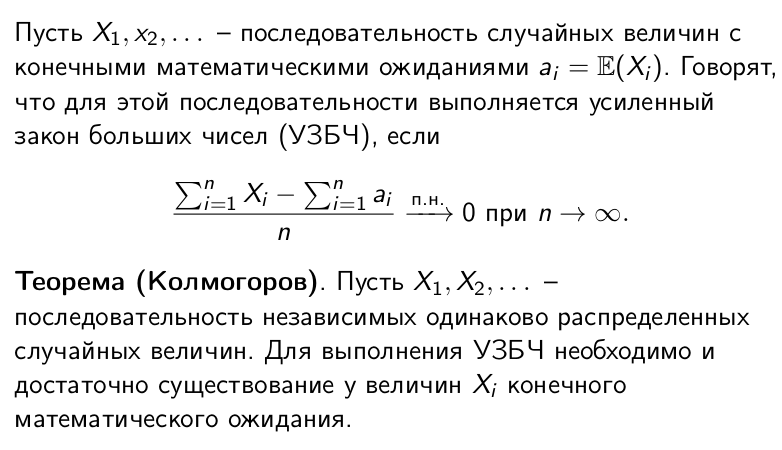
**Đặt X1 , x2 , . . . – dãy các biến ngẫu nhiên có kỳ vọng toán học hữu hạn ai = E(Xi ). Luật số lớn được cho là đúng cho dãy này nếu**

****

**Định lý (Khichin). Nếu X1 , X2 , . . . là một chuỗi các biến ngẫu nhiên được phân phối giống hệt độc lập với kỳ vọng toán học hữu hạn thì luật số lớn được thỏa mãn.**

1. **Усиленный закон больших чисел - Luật số lớn được tăng cường**

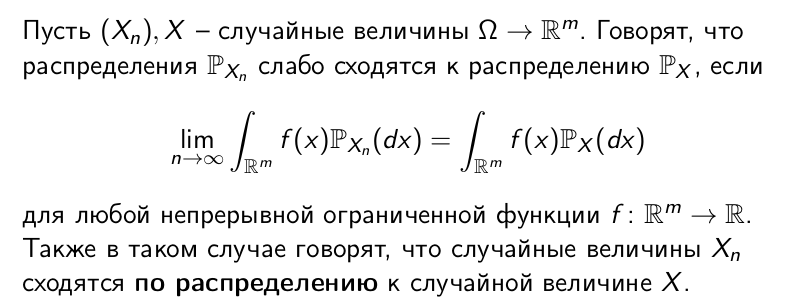
**Đặt X1 , x2 , . . . – dãy các biến ngẫu nhiên có kỳ vọng toán học hữu hạn ai = E(Xi ). Chuỗi này được cho là thỏa mãn Luật mạnh về số lớn nếu**

****

**Định lý (Kolmogorov). Đặt X1 , X2 , . . . – một chuỗi các biến ngẫu nhiên được phân phối giống hệt nhau độc lập. Để đáp ứng luật số lớn tăng cường, điều cần và đủ là các giá trị Xi có kỳ vọng toán học hữu hạn.**

1. **Слабая сходимость распределений - Sự hội tụ yếu của phân phối**

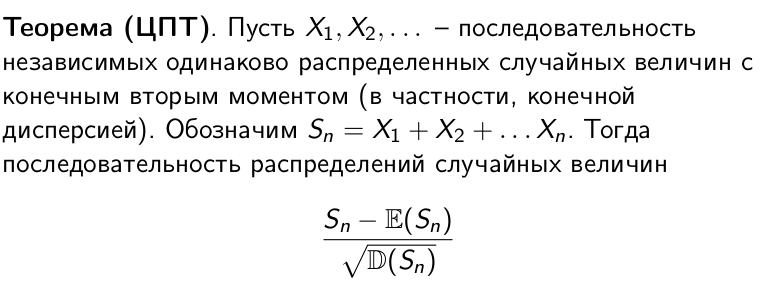
**cho (Xn), X – biến ngẫu nhiên Ω → Rm. Phân bố PXn được cho là hội tụ yếu về phân bố PX nếu**

****

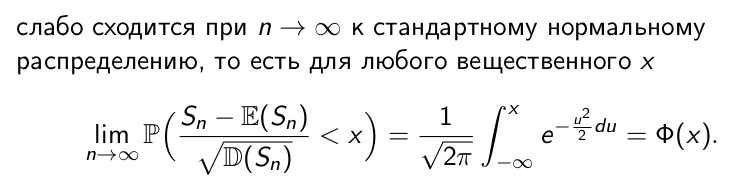
**đối với hàm giới hạn liên tục bất kỳ f: Rm → R. Cũng trong trường hợp này, chúng ta nói rằng các biến ngẫu nhiên Xn hội tụ theo phân phối cho biến ngẫu nhiên X.**

1. **Центральная предельная теорема - Định lý giới hạn trung tâm**

**Định lý (CPT). Đặt X1 , X2 , . . . – một chuỗi các biến ngẫu nhiên được phân phối giống hệt độc lập với mô men thứ hai hữu hạn (đặc biệt là phương sai hữu hạn). Ta ký hiệu Sn = X1 + X2 + . . . Xn. Khi đó trình tự phân phối của các biến ngẫu nhiên**

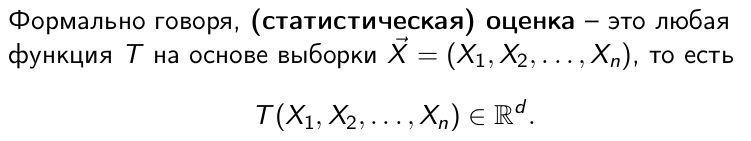
****

**hội tụ yếu khi n → ∞ theo phân phối chuẩn chuẩn hóa, nghĩa là với mọi x thực**

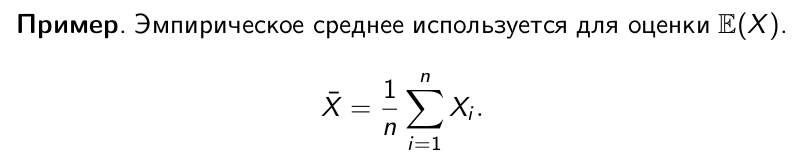
****

1. **Статистическая оценка - Đánh giá thống kê**

**Nói một cách chính thức, ước tính (thống kê) là bất kỳ hàm T nào dựa trên mẫu X= (X1 , X2 , . . . , Xn ), nghĩa là**

****

**Ví dụ. Giá trị trung bình thực nghiệm được sử dụng để ước tính E(X).**

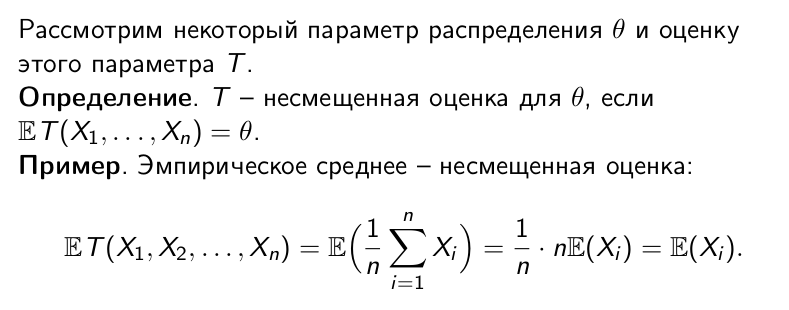
****

1. **Несмещенные оценки**

**Hãy xem xét một số tham số phân phối θ và ước tính tham số T này.**

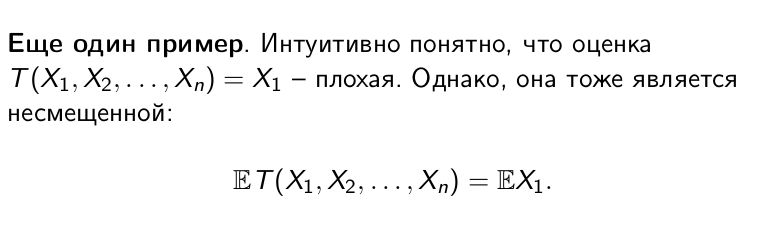
**Sự định nghĩa. T là ước lượng không chệch cho θ nếu ET (X1 , . . . , Xn ) = θ.**

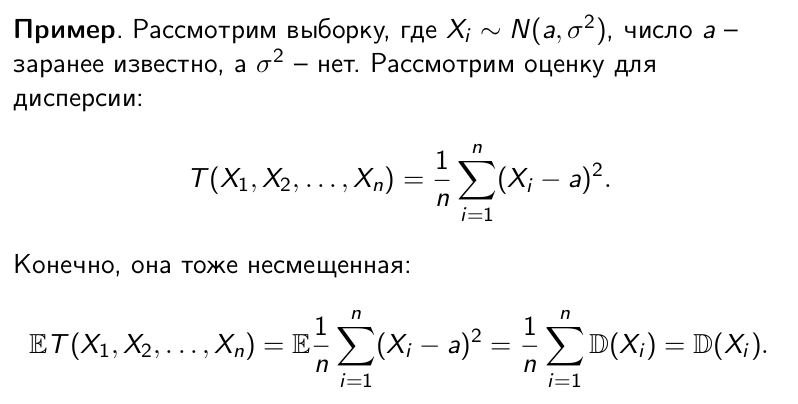
**Ví dụ. Ý nghĩa thực nghiệm - ước tính không thiên vị:**

****

**Một ví dụ nữa. Bằng trực giác, rõ ràng rằng ước lượng T (X1, X2, . . ., Xn) = X1 là sai. Tuy nhiên, nó cũng không thiên vị:**

**ET(X1, X2, . . , Xn) = EX1.**

****

****

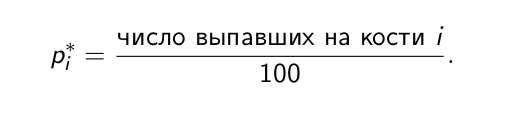
1. **Задачи математической статистики**

**Các vấn đề về thống kê toán học**

* Предположим, что мы повторяем один и тот же случайный эксперимент в одинаковых условиях. Что можно о распределении в практическом эксперименте? Пример. С какой вероятностью выпадает герб на данной монете?
* **Giả sử chúng ta lặp lại cùng một thí nghiệm ngẫu nhiên trong cùng điều kiện. Bạn có thể nói gì về sự phân bố trong một thí nghiệm thực tế? Ví dụ. Xác suất để có được huy hiệu trên một đồng tiền nhất định là bao nhiêu?**
* Предположим, что мы повторяем один и тот же случайный эксперимент в одинаковых условиях. Что можно о распределении в практическом эксперименте? Пример. С какой вероятностью выпадает герб на данной монете? Для определения вероятности мы можем подбросить монету много раз, но выводы придется сделать по результатам конечного числа наблюдений. Если после 10000 бросков монеты выпадет 5035 гербов, нельзя сделать точный вывод о вероятности выпадения герба.
* **Giả sử chúng ta lặp lại cùng một thí nghiệm ngẫu nhiên trong cùng điều kiện. Bạn có thể nói gì về sự phân bố trong một thí nghiệm thực tế? Ví dụ. Xác suất để có được huy hiệu trên một đồng tiền nhất định là bao nhiêu? Chúng ta có thể tung đồng xu nhiều lần để xác định xác suất, nhưng chúng ta phải đưa ra kết luận dựa trên số lượng quan sát hữu hạn. Nếu sau 10.000 lần tung đồng xu, 5.035 biểu tượng xuất hiện thì không thể rút ra kết luận chính xác nào về khả năng quốc huy bị loại bỏ.**
* Пусть ξ : Ω → R – случайная величина, наблюдаемая в случайном эксперименте. Пусть X1 , X2 , . . . , Xn – значения наблюдаемой случайной величины. Случайная величина ξ имеет некоторое распределение F, которое нам частично или полностью неизвестно.
* **Đặt ξ : Ω → R là biến ngẫu nhiên được quan sát trong một thí nghiệm ngẫu nhiên. Đặt X1 , X2 , . . . , Xn – giá trị của biến ngẫu nhiên quan sát được. Biến ngẫu nhiên ξ có phân phối F nào đó mà chúng ta chưa biết một phần hoặc hoàn toàn.**
* Удобно считать, что до опыта Xi – случайная величина, одинаково распределенная с ξ, а после опыта Xi – число, которое мы наблюдаем в i-м по счету эксперименте, то есть одно из возможных значений случайной величины Xi . ⃗ = (X1 , . . . , Xn ) объема n из Определение. Выборкой X распределения F называется набор из n независимых и одинаково распределенных случайных величин, имеющих распределение F.
* **Thật thuận tiện khi giả sử rằng trước thí nghiệm Xi là một biến ngẫu nhiên được phân phối giống hệt với ξ và sau thí nghiệm Xi là số mà chúng ta quan sát được trong thí nghiệm thứ i, nghĩa là một trong các giá trị có thể có của ngẫu nhiên biến Xi. ⃗ = (X1, . . . , Xn) của tập n từ Định nghĩa. Mẫu X của phân phối F là tập hợp n biến ngẫu nhiên độc lập và được phân phối giống hệt nhau có phân phối F.**
* Задачей математической статистики является получение по выборке выводов о неизвестном распределении F, из которого она извлечена. Распределение характеризуется функцией распределения, плотностью или набором числовых характеристик. По выборке строятся приближения для этих характеристик, так называемые оценки.
* **Nhiệm vụ của thống kê toán học là rút ra kết luận từ một mẫu về phân bố F chưa biết mà từ đó nó được rút ra. Một phân phối được đặc trưng bởi hàm phân phối, mật độ hoặc một tập hợp các đặc tính số. Dựa trên mẫu, các giá trị gần đúng cho những đặc điểm này, được gọi là ước tính, được xây dựng.**

Пример. Шестигранный кубик подброшен 100 раз. Первая грань выпала 25 раз, вторая и пятая – по 14 раз, третья – 21 раз, четвертая – 15 раз, шестая – 11 раз. Оценкой для неизвестной вероятности pi (i = 1, 2, 3, 4, 5, 6) будет случайная величина

**Ví dụ. Xúc xắc sáu mặt được tung 100 lần. Mặt thứ nhất rơi ra 25 lần, mặt thứ hai và thứ năm – mỗi mặt 14 lần, mặt thứ ba – 21 lần, mặt thứ tư – 15 lần, mặt thứ sáu – 11 lần. Ước tính cho xác suất chưa biết pi (i = 1, 2, 3, 4, 5, 6) sẽ là một biến ngẫu nhiên**

****

# **Лекция 2. Выборка. Эмпирическая функция. Выборочные моменты**

1. **Статистическая модель - Mô hình thống kê**

* X1 , X2 , . . . , Xn – выборка (n – объем выборки);
* P = {P} – семейство возможных распределений. P = {P} – họ các phân bố có thể có.

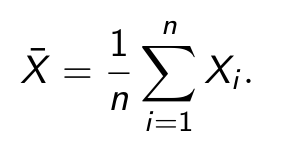
Def.

Статистика – любая функция от выборки.

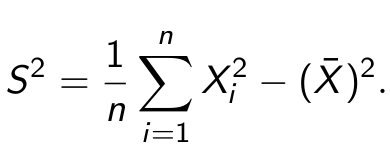
**Thống kê là bất kỳ chức năng nào của một mẫu.**

1. **Характеристики выборки**

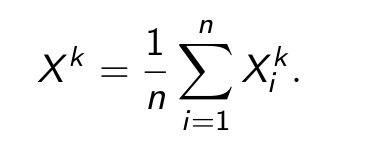
* **Эмпирическое среднее:** Trung bình thực nghiệm:



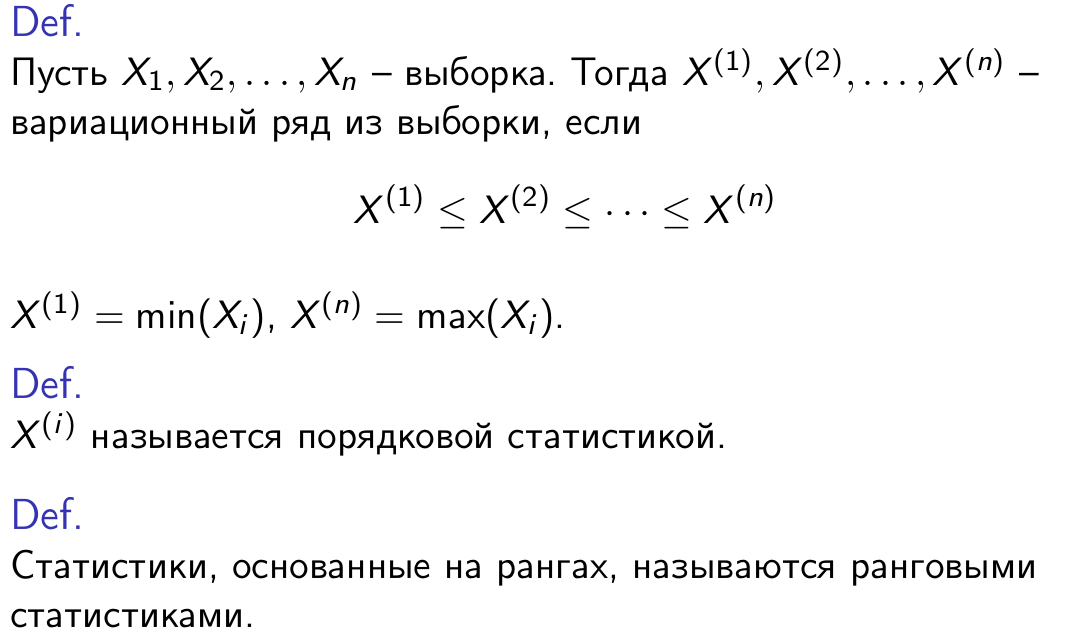
* **Эмпирическая дисперсия: Phương sai**

****

* **Выборочный момент порядка k:**

****

1. **Вариационный ряд: chuỗi biến thể**

****

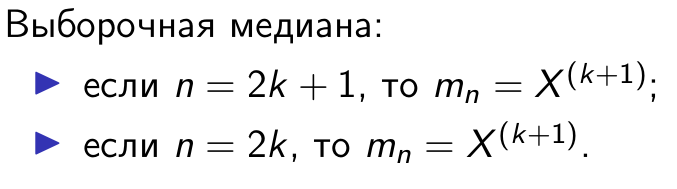
Đặt X1 , X2 , . . . , Xn – mẫu. Khi đó X (1) , X (2) , . . . , X(n) – chuỗi biến thiên từ mẫu, nếu

X(i) được gọi là thống kê thứ tự.

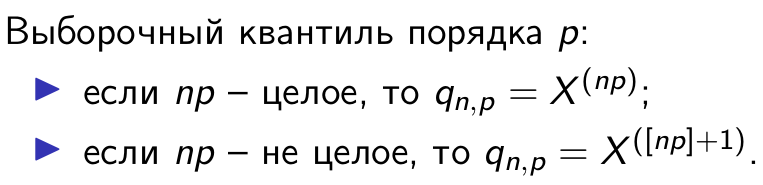
Thống kê dựa trên thứ hạng được gọi là thống kê thứ hạng.

1. **Характеристики выборки:**

**Trung vị mẫu:**

****

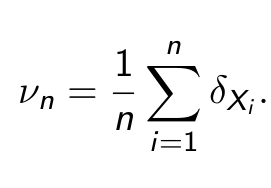
**Lượng tử mẫu theo thứ tự p:**

****

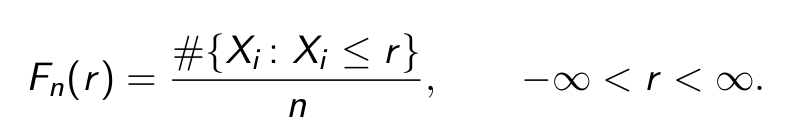
1. **Эмпирическая функция распределения**

**Hàm phân phối theo kinh nghiệm**

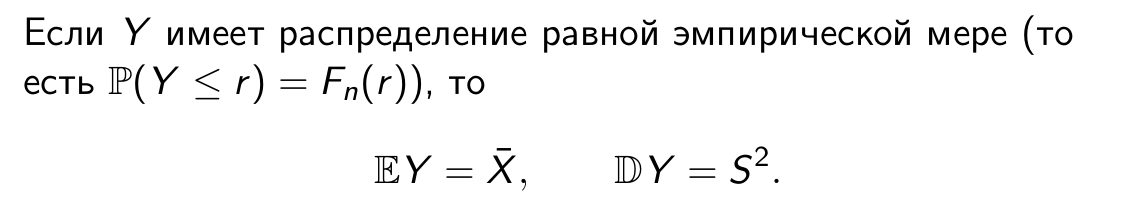
* Эмпирическая мера: Biện pháp thực nghiệm



* Эмпирическая функция распределения:

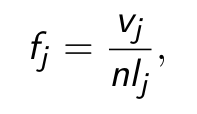


* Nếu Y có phân bố bằng thước đo thực nghiệm (nghĩa là P(Y ≤ r) = Fn(r)), thì



1. **Гистограмма: Biểu đồ cột**

* **Эмпирическим аналогом плотности распределения является так называемая гистограмма. Гистограмма строится по группированным данным. Область на прямой, занимаемую элементами выборки, делят на k интервалов. Пусть A1 , . . . , Ak — интервалы на прямой, называемые интервалами группировки . Обозначим для j = 1, . . . , k через vj число элементов выборки, попавших в интервал Aj .**
* Một sự tương tự thực nghiệm của mật độ phân bố được gọi là biểu đồ. Biểu đồ được xây dựng bằng cách sử dụng dữ liệu được nhóm. Diện tích trên đường thẳng chứa các phần tử mẫu được chia thành k khoảng. Cho A1 , . . . , Ak - các khoảng trên đường thẳng, gọi là các khoảng nhóm. Chúng ta ký hiệu cho j = 1, . . . , k đến vj là số phần tử mẫu rơi vào khoảng Aj.
* **На каждом из интервалов Aj строят прямоугольник, площадь которого пропорциональна vj . Общая площадь всех прямоугольников должна равняться единице. Поэтому высота fj прямоугольника над интервалом Aj равна**
* Tại mỗi khoảng Aj, một hình chữ nhật được dựng lên, diện tích của nó tỉ lệ với vj. Tổng diện tích của tất cả các hình chữ nhật phải bằng một. Do đó, chiều cao fj của hình chữ nhật trên đoạn Aj bằng



**где через lj обозначена длина интервала Aj . Полученная фигура, состоящая из объединения прямоугольников, называется гистограммой.**

trong đó lj biểu thị độ dài của khoảng Aj. Hình kết quả, bao gồm các hình chữ nhật kết hợp, được gọi là biểu đồ.

**Пример. Имеется вариационный ряд: Ví dụ. Có một phạm vi biến thể:**

(0; 1; 1; 2; 2, 6; 2, 6; 2, 6; 3, 1; 4, 6; 4, 6; 6; 6; 7; 9; 9).

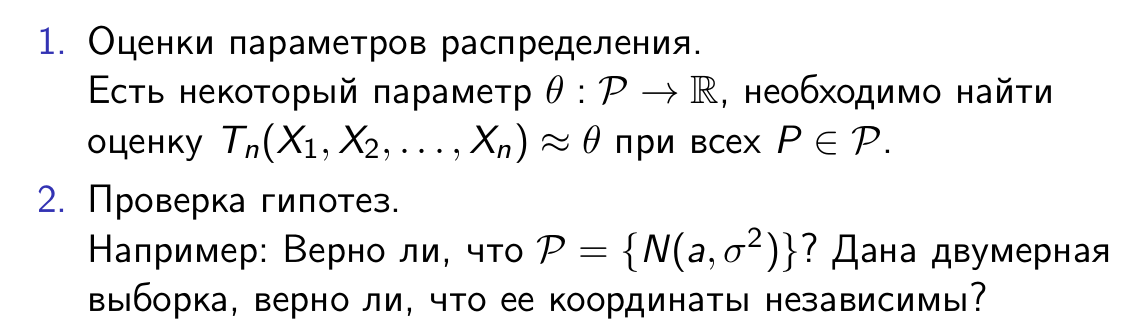
**Разобьём отрезок [0, 10] на четыре равных отрезка. Отрезку [0, 2, 5) принадлежат четыре элемента выборки, отрезку [2, 5, 5) — шесть, отрезку [5, 7, 5) — три, и отрезку [7, 5, 10] — два элемента выборки.**

**Hãy chia đoạn [0, 10] thành bốn đoạn bằng nhau. Đoạn [0, 2, 5) có bốn phần tử mẫu, đoạn [2, 5, 5) có sáu, đoạn [5, 7, 5) có ba và đoạn [7, 5, 10] có hai các phần tử mẫu.**

1. **Классы задач мат. статистики - Các lớp giải toán. số liệu thống kê**

**1. Ước tính các tham số phân phối. Có tham số θ : P → R, cần tìm ước lượng Tn (X1, X2, . . . , Xn) ≈ θ với mọi P ∈ P.**

**2. Kiểm tra giả thuyết. Ví dụ: Có đúng P = {N(a, σ ^2 )} không? Cho một mẫu hai chiều, liệu tọa độ của nó có độc lập không?**

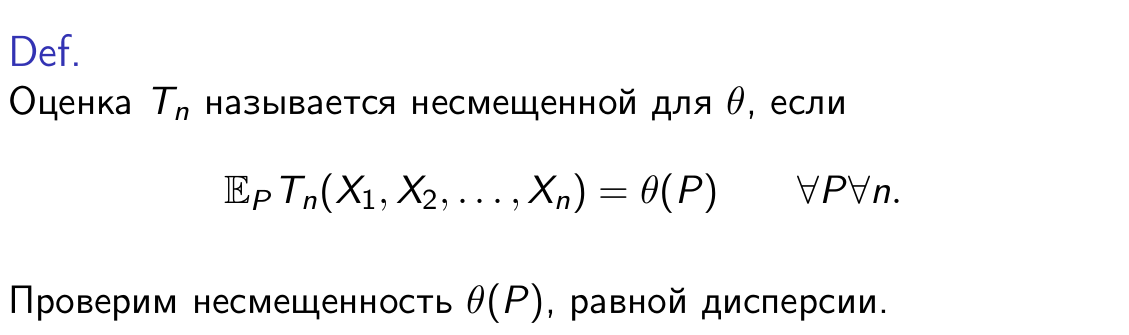
****

1. **Несмещенные оценки - ước tính không thiên vị**

Ước lượng Tn được gọi là không chệch cho θ nếu

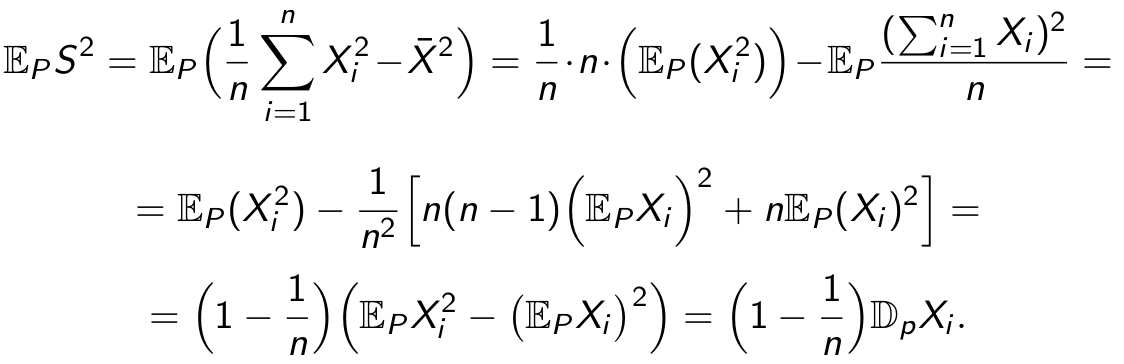
EP Tn (X1 , X2 , . . . , Xn ) = θ(P) ∀P∀n.

Chúng ta hãy kiểm tra xem θ(P) có không thiên vị và bằng phương sai hay không.



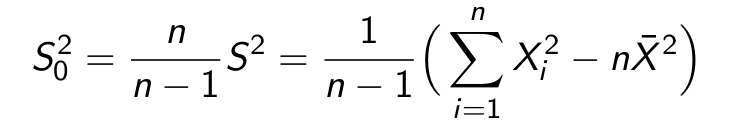
**Проверим несмещенность θ(P), равной дисперсии.**

Chúng ta hãy kiểm tra xem θ(P) có không thiên vị và bằng phương sai hay không.



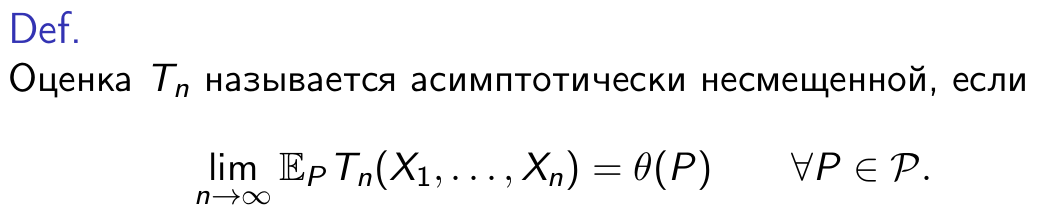
Таким образом, S^2 – смещенная оценка для дисперсии

Do đó, S^2 là ước tính sai lệch cho phương sai



несмещенная оценка для дисперсии. ước lượng không chệch cho phương sai.

Ước lượng Tn được gọi là không thiên lệch tiệm cận nếu

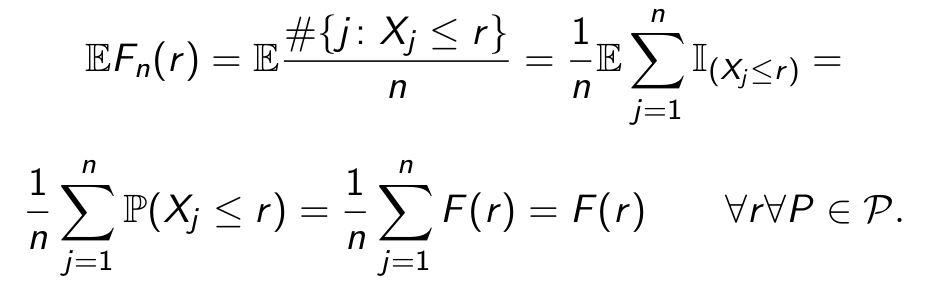


**Утверждение**

Эмпирическая функция распределения является несмещенной оценкой для истинной функции распределения F .

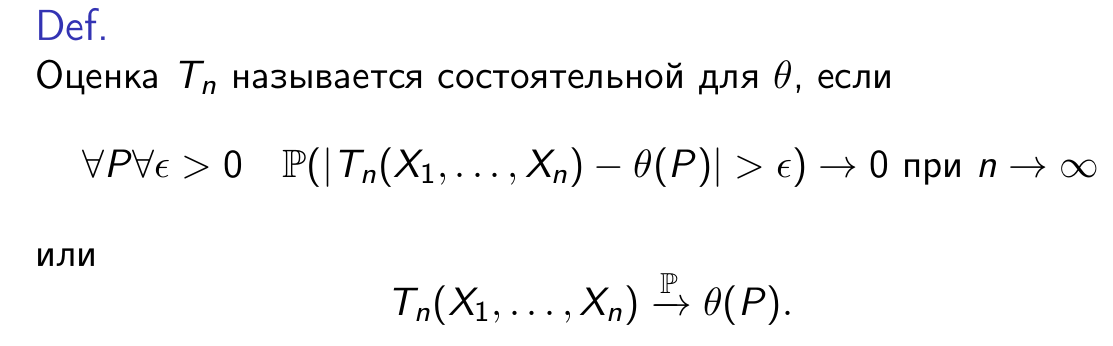
**Hàm phân phối thực nghiệm là ước tính không chệch của hàm phân phối thực F.**

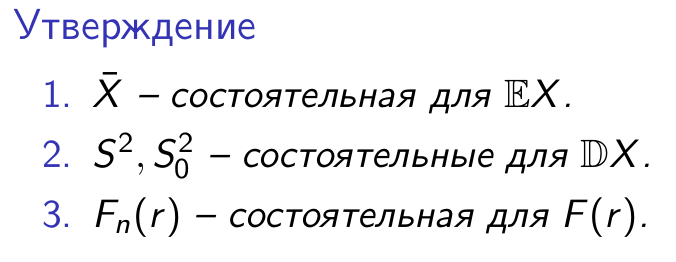
**Доказательство**.

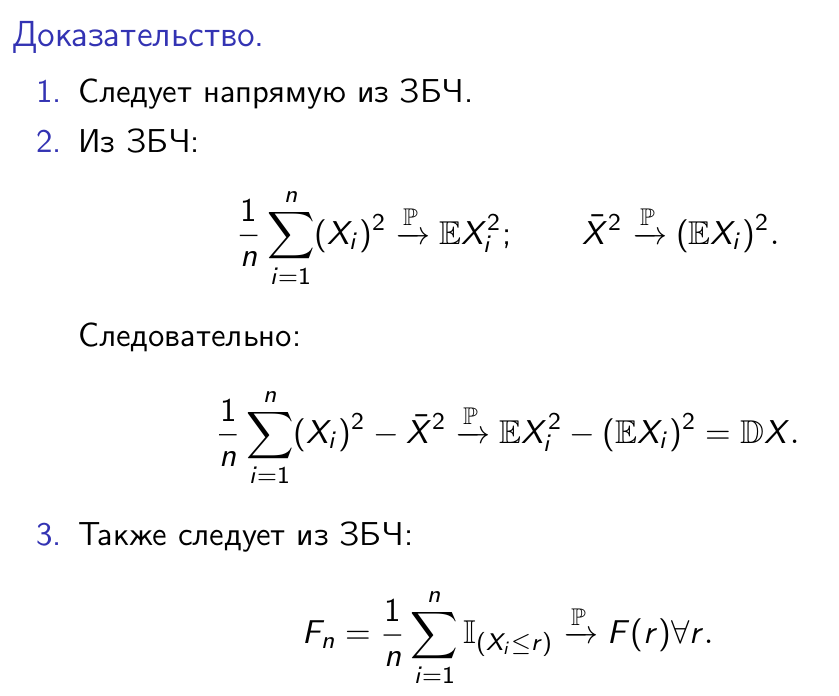


1. **Состоятельность**

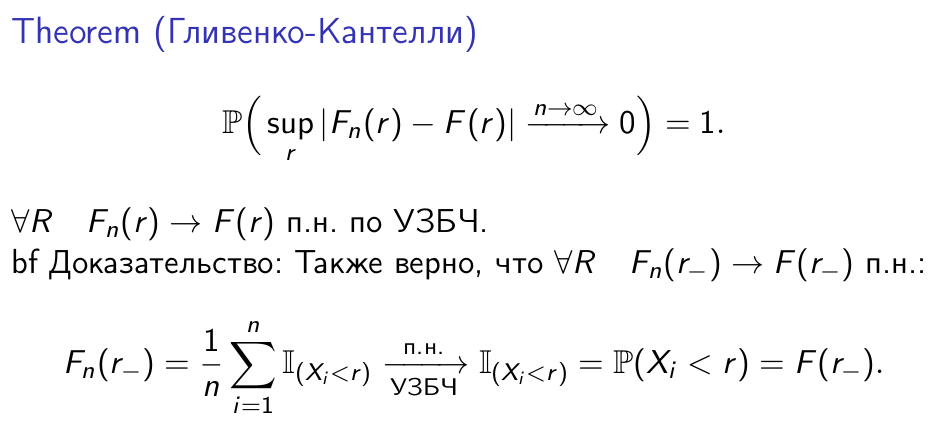
**Ước lượng Tn được cho là nhất quán với θ nếu**

****

****

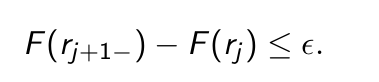
****

1. **Теорема Гливенко-Кантелли - Định lý Glivenko-Cantelli**

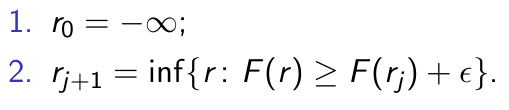
****

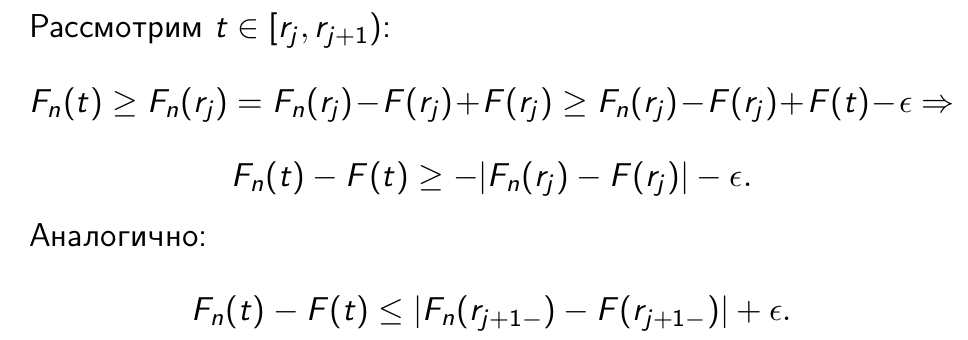
**Зафиксируем ϵ > 0. Построим последовательность r0 = −∞ < r1 < · · · < rm = ∞ со следующими свойствами:**

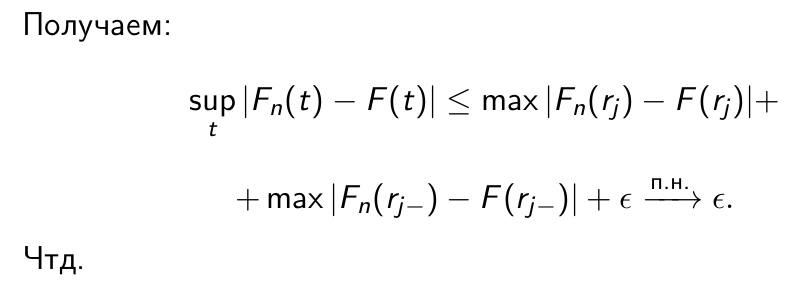
Hãy sửa ϵ > 0. Xây dựng một dãy r0 = −∞ < r1 < · · · < rm = ∞ với các tính chất sau:



Можно построить по индукции: Ta có thể xây dựng bằng quy nạp:







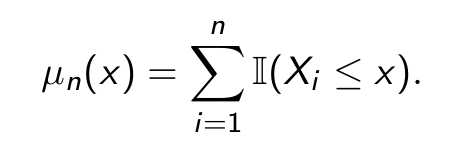
# **Лекция 3. Различные асимптотические свойства**

Bài 3. Các tính chất tiệm cận khác nhau

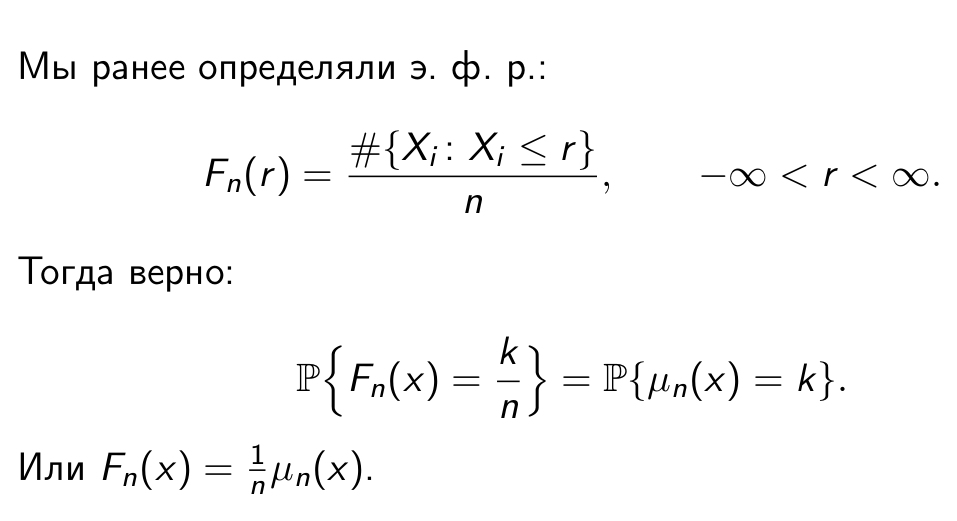
1. **Обозначения µn (x) - Ký hiệu µn(x)**

Определим для каждого действительного числа x случайную величину µn (x), равную числу элементов выборки X = (X1 , . . . , Xn ), значения которых не превосходят x, то есть

Với mỗi số thực x, ta xác định một biến ngẫu nhiên µn(x) bằng số phần tử mẫu X = (X1 , . . . , Xn ), các giá trị của chúng không vượt quá x, tức là



1. **Эмпирическая функция распределения - Hàm phân phối theo kinh nghiệm**

****