Случайные события

1. Что понимают под элементарным исходом (элементарным событием)? Дайте определение вероятности по Лапласу (классическое) и укажите на его недостатки. Сформулируйте и докажите свойства вероятности (для вероятности по Лапласу).

Элементарный исход – любой простейший исход испытания

Вероятностью события A называется отношение числа N_A благоприятствующих событию A элементарных исходов к общему числу N элементарных исходов в Ω

$P = \frac{N_A}{N}$

Недостатки: требование конечности пространства элементарных исходов и равновозможности элементарных исходов

Свойства:

1)$\forall$ A $\subset$ Ω $P(A) \geq 0$

2) P(Ω) = 1

3) P(A+B) = P(A) + P(B), если A и B несовместные

Доказательство:

1) очевидно, так как N_A и N ≥ 0

2)$P(\Omega) = \frac{N}{N} = 1$

3) событию A благоприятствует N_A исходов, а событию B – N_B исходов. A и B несовместные, следовательно, событию A+B благоприятствует $N_A + N_B$ исходов. Тогда

$$P(A+B) = \frac{N_A + N_B}{N} = \frac{N_A}{N} + \frac{N_B}{N} = P(A) + P(B)$$

2. Элементарный исход и случайное событие. Дать определение операций, определённых для случайных событий.

   Элементарный исход – любой простейший исход испытания
   Случайное событие – испытание, наблюдаемые результаты которого случайны.
   Пересечением двух событий A и B называется событие

$C = A \cap B = AB = \{\omega \in \Omega | \omega \in A \text{ и } \omega \in B\}$

Объединением двух событий A и B называется событие

$C = A \cup B = A + B = \{\omega \in \Omega | \omega \in A \text{ или } \omega \in B\}$

Разностью двух событий A и B называется событие

$C = A \setminus B = AB = \{\omega \in \Omega | \omega \in A \text{ и } \omega \notin B\}$

Дополнением события A называется событие $\overline{A} = \{\omega \in \Omega | \varpi \notin A\}$

3. Элементарный исход, пространство элементарных исходов. Дайте геометрическое определение вероятности. Какими специфическими особенностями обладает геометрическая вероятность?

Элементарный исход – любой простейший исход испытания

Множество всех элементарных исходов образует пространство элементарных исходов Ω, если выполнено:

1) В результате опыта один из исходов обязательно происходит
2) Появление одного из исходов опыта исключает появление остальных
3) В рамках данного опыта нельзя разделить элементарные исходы на более мелкие

Вероятность события $A \subset \Omega$ есть число $P(A) = \frac{\mu(A)}{\mu(\Omega)}$, где μ(A) и μ(Ω) – меры множеств A и Ω соответственно

Особенности:

1)Если размерность множества A меньше размерности множества Ω, то μ(A) = 0 и P(A) = 0 - парадокс нулевой вероятности

2) требуется равновозможность элементарных исходов

4. Элементарный исход, пространство элементарных исходов. Дайте определение вероятности по Колмогорову (аксиоматическое).

Элементарный исход – любой простейший исход испытания

Множество всех элементарных исходов образует пространство элементарных исходов Ω, если выполнено:

1) В результате опыта один из исходов обязательно происходит
2) Появление одного из исходов опыта исключает появление остальных

3) В рамках данного опыта нельзя разделить элементарные исходы на более мелкие

Пусть каждому событию А поставлено в соответствие число P(A). Числовую функцию P, заданную на σ-алгебре B, называют вероятностью или вероятностной мерой, если она удовлетворяет трем аксиомам:

1) $P(A) \geq 0$
2) $P(\Omega)=1$
3) $\forall$ попарно несовместных событий $A_1, \ldots, A_n, \ldots$
   $P(A_1, \ldots, A_n, \ldots) = P(A_1)+\ldots+P(A_n)+\ldots$

Значение P(A) называют вероятностью события А

5. Элементарный исход, пространство элементарных исходов. Дать определение попарно несовместных и несовместных в совокупности событий.

Элементарный исход – любой простейший исход испытания

Множество всех элементарных исходов образует пространство элементарных исходов Ω, если выполнено:

1) В результате опыта один из исходов обязательно происходит
2) Появление одного из исходов опыта исключает появление остальных
3) В рамках данного опыта нельзя разделить элементарные исходы на более мелкие

События $A_1, \ldots, A_n$ называются попарно несовместными тогда и только тогда, когда $A_i A_j = \emptyset \ \forall \ i, j = 1, \ldots, n$ и $i \neq j$

События $A_1, \ldots, A_n$ называются несовместными в совокупности тогда и только тогда, когда $A_1 A_2 \ldots A_n = \emptyset$

6. Элементарный исход, пространство элементарных исходов. Дать определение попарно независимых и независимых в совокупности событий.

Элементарный исход – любой простейший исход испытания

Множество всех элементарных исходов образует пространство элементарных исходов Ω, если выполнено:

1) В результате опыта один из исходов обязательно происходит
2) Появление одного из исходов опыта исключает появление остальных
3) В рамках данного опыта нельзя разделить элементарные исходы на более мелкие

События $A_1,\dots,A_n$ называются попарно независимыми тогда и только тогда, когда $P(A_iA_j)=P(A_i)P(A_j)\ \forall\ i,j=1,\dots,n$ и $i\neq j$

События $A_1,\dots,A_n$ называются независимыми в совокупности тогда и только тогда, когда $P(A_1A_2\dots A_n)=P(A_1)P(A_1)\dots P(A_n)$

7. Дать определение несовместных и независимых случайных событий. Как связаны между собой несовместные и независимые случайные события?

События A и B называют несовместными, если $AB=\emptyset$

События A и B, имеющие ненулевую вероятность, называют независимыми, если их условные вероятности совпадают с безусловными, то есть $P(A|B)=P(A)$ и $P(\mathrm{B}|\mathrm{A})=P(\mathrm{B})$

Связь:

1) Если A и B несовместные и $P(A)\neq 0$, $P(B)\neq 0$, то они зависимые
2) Если A и B совместные, то они могут быть и зависимые, и независимые
3) Если A и B зависимые, то они могут быть и совместными, и несовместными

8. Дать определение σ-алгебры, σ-алгебры событий, алгебры событий.

σ-алгеброй B называют непустую систему подмножеств U, удовлетворяющую следующим условиям:

1) Если подмножество $A\in B$, то и $\overline{A}\in B$
2) Если подмножества $A_1,A_2,\dots,A_n,\dots\in B$, то $A_1+A_2+\dots+A_n+\dots\in B$ и $A_1A_2\dots A_n\dots\in B$

Элементы σ-алгебры B, заданной на пространстве элементарных исходов **Ω**, называют событиями, а саму σ-алгебру B называют σ-алгеброй событий

Если в условии 2 счетное множество событий заменить на конечное, то получим определение алгебры событий

9. Дать определение полной группы событий, гипотезы. Сформулировать и доказать теорему о формуле полной вероятности.

Совокупность случайных событий $H_1,H_2,\dots,H_n$ называют гипотезами, если:

1) $H_i \neq \emptyset \ \forall \ i=1,\dots,n$
2) $Hi \cap Hj = \emptyset \ \forall \ i\neq j$
3) $\bigcup_{i=1}^{n} H_i = \Omega$

Если условие 2 не выполнено, то события $H_1,H_2,\dots,H_n$ называют полной группой событий

Теорема:

Пусть событие A и гипотезы $H_1,\dots,H_n$ заданы на одном и том же вероятностном пространстве и известны $P(H_i)$ и $P(A|H_i) \ \forall \ i=1,\dots,n$. Тогда вероятность P(A) можно определить по формуле $P(A)=P(H_1)P(A|H_1)\dots P(A)P(A|H_n)$

Доказательство:

Так как $A=A\Omega=A(H_1+\dots+H_n)=AH_1+\dots+AH_n$ и так как $(AH_i) \cap \left(AH_j\right) = \emptyset \ \forall \ i \neq j$, то $P(A)=P(\bigcup_{i=1}^{n}(AH_i))=\sum_{i=1}^{n} P(AH_i) = \sum_{i=1}^{n} P(H_i)P(A|H_i)$

10. Дать определение полной группы событий, гипотезы. Сформулировать и доказать теорему Байеса.

Совокупность случайных событий $H_1,H_2,\dots,H_n$ называют гипотезами, если:

1) $H_i \neq \emptyset \ \forall \ i=1,\dots,n$
2) $Hi \cap Hj = \emptyset \ \forall \ i\neq j$

3) $\bigcup_{i=1}^{n} H_i = \Omega$

Если условие 2 не выполнено, то события $H_1, H_2, \ldots, H_n$ называют полной группой событий

Теорема:

Пусть для некоторого события A $P(A) \geq 0$ и для гипотез $H_1, \ldots, H_n$ известны $P(H_i) \geq 0$ и $P(A|H_i)$ $i,j=1,\ldots,n$. Тогда $P(H_i|A)=\frac{P(H_i)P(A|H_i)}{\sum_1^n P(H_i)P(A|H_i)}$

Доказательство:

По определению $P(H_i|A) = \frac{P(AH_i)}{P(A)}$

По формуле умножения вероятностей $P(AH_i)=P(H_i)P(A|H_i)$

По формуле полной вероятности $P(A)=\sum_1^n P(A|H_i)P(H_i)$

Тогда $P(H_i|A)= \frac{P(H_i)P(A|H_i)}{\sum_1^n P(H_i)P(A|H_i)}$

11. Дать определение условной вероятности. Сформулировать и доказать теорему умножения.

Условной вероятностью события A при условии наступления события B называют следующее число: $P(A|B) = \frac{P(AB)}{P(B)}$, при $P(B) \neq 0$

Теорема:

Пусть событие $A=A_1A_2\ldots A_n$ и $P(A) > 0$. Тогда справедливо равенство $P(A_1A_2\ldots A_n)=P(A_1)P(A_2|A_1)\ldots P(A_n|A_1A_2\ldots A_{n-1})$

Доказательство:

Так как $P(A)=P(A_1A_2\ldots A_n)>0$ и $A_1A_2\ldots A_n \subset A_1A_2\ldots A_k$ $\forall\, k = 1,\ldots,n-1$, то $P(A_1A_2\ldots A_k)>0$

$P(A_n| A_1A_2\ldots A_{n-1})=\frac{P(A_1A_2\ldots.A_n)}{A_1A_2\ldots A_{n-1}}$

следовательно $P(A_1A_2\ldots A_n)=P(A_1A_2\ldots A_{n-1})\, P(A_n| A_1A_2\ldots A_{n-1})$

Аналогично $P(A_1A_2\ldots A_{n-1})=P(A_1A_2\ldots A_{n-2})\, P(A_{n-1}| A_1A_2\ldots A_{n-2})$

Тогда $P(A_1A_2\ldots A_n)=P(A_1A_2\ldots A_{n-2})\ P(A_{n-1}|\ A_1A_2\ldots A_{n-2})\ P(A_n|\ A_1A_2\ldots A_{n-1})$ и так далее

12. Что понимают под биномиальной схемой испытаний (схемой Бернулли)? Сформулировать и доказать теорему Бернулли.

Биномиальной схемой испытаний (Схемой Бернулли) называют последовательность повторных испытаний, удовлетворяющих условиям:

1) Для каждого испытания возможно лишь два исхода: появления некоторого события A («успех») либо его дополнения $\overline{A}$ («неудача»)
2) Исход любого испытания не зависит от исходов предшествующих испытаний, то есть испытания независимы
3) Вероятность успеха во всех испытаниях постоянна и равна P(A)

Теорема:

Вероятность $P_n(k)$ того, что в n испытаниях по схеме Бернулли произойдет ровно k «успехов» равна $P_n(k)=C_n^k\ p^k\ q^{n-k}$, где $C_n^k=\frac{n!}{(n-k)!k!}$, $k=0,\ldots,n$

Доказательство:

Пусть событие B_h состоит в том, что в испытаниях с номерами $i_1^h, i_2^h, \ldots, i_k^h$ наступило событие «успех», а с номерами $j_1^h, j_2^h, \ldots, j_{n-k}^h$ – «неудача»

Тогда $\{i_1^h, i_2^h, \ldots, i_k^h\} \cup \{j_1^h, j_2^h, \ldots, j_{n-k}^h\}=\{1,2,\ldots,n\}$, а

$\{i_1^h, i_2^h, \ldots, i_k^h\} \cap \{j_1^h, j_2^h, \ldots, j_{n-k}^h\} = \emptyset$

Таким образом, $B_h^k \Leftrightarrow \left\{(A_{i_1{}^h}) \wedge \ldots \wedge (A_{i_k{}^h})\right\} \wedge \left\{(\overline{A_{J_1{}^h}}) \wedge \ldots \wedge (\overline{A_{J_{n-k}{}^h}})\right\}$

Поскольку события независимые:

$P(B_h)=P(A_{i_1{}^h})\ldots\ P(A_{i_k{}^h})\ P(\overline{A_{J_{n-k}{}^h}})= p^k(1-p)^{n-k}$

Тогда событие $B_n^k=\bigcup_h B_h$ – в n испытаниях k «успехов», где h – количество способов выбрать k «успешных» мест в n испытаниях. Это число равно $C_n^k = \frac{n!}{(n-k)!k!}$

Тогда $P(B_h^k)=P(\bigcup_h B_h) = \sum_h P(B_h) = C_n^k p^k (1-p)^{n-k}$

13. Что понимают под биномиальной схемой испытаний (схемой Бернулли)? Сформулировать следствия из теоремы Бернулли.

Биномиальной схемой испытаний (Схемой Бернулли) называют последовательность повторных испытаний, удовлетворяющих условиям:

1) Для каждого испытания возможно лишь два исхода: появления некоторого события A («успех») либо его дополнения $\overline{A}$ («неудача»)
2) Исход любого испытания не зависит от исходов предшествующих испытаний, то есть испытания независимы
3) Вероятность успеха во всех испытаниях постоянна и равна P(A)

Теорема:

Вероятность $P_n(k)$ того, что в n испытаниях по схеме Бернулли произойдет ровно k «успехов» равна $P_n(k)=C_n^k p^k q^{n-k}$, где $C_n^k=\frac{n!}{(n-k)!k!}$, k=0,…,n

Следствия:

1) $P\{K_1 \le K \le K_2\} = \sum_{K_1}^{K_2} C_n^k p^k q^{n-k}$
2) $P\{K \ge 1\} = 1 - q^n$

14. Дать определение несовместных событий. Вывести формулу сложения вероятностей для 2-х и для n событий.

События A и B называют несовместными, если $AB=\emptyset$

Для 2 событий:

P(A+B)=P(A)+P(B)-P(AB)

Доказательство:

$\begin{cases} A + B = A + (B|A) \\ B = (B|A) + AB \end{cases} \Rightarrow \begin{cases} P(A + B) = P(A) + P(B|A) \\ P(B) = P(B|A) + P(AB) \end{cases} \Rightarrow P(A + B) = P(A) + P(B) - P(AB)$

Для n событий:

$P(A_1+\ldots+A_n)=P(A_1)+\ldots+P(A_n) - P(A_1A_2) - P(A_1A_3) - \ldots - P(A_{n-1}A_n) + P(A_1A_2A_3)+\ldots+(-1)^{n+1}P(A_1A_2\ldots A_n)$

Доказательство на примере трех:

P(A+B+C)=P(A)+P(B+C) – P(A|B+C)=P(A) + P(B) + P(C) – P(BC) – P(AB + AC) = P(A) + P(B) + P(C) – P(BC) – P(AB) – P(AC) + P(ABC)

15. Сформулировать и доказать основные свойства вероятности (6 свойств).

1) $P(\bar{A})=1 - P(A)$
2) $P(\emptyset)= 0$
3) $A \subset B \Rightarrow P(A) \leq P(B)$
4) $0 \leq P(A) \leq 1$
5) $P(A+B)=P(A)+P(B)-P(AB)$
6) $P(A_1+\ldots+A_n)=P(A_1)+\ldots+P(A_n) - P(A_1A_2) - P(A_1A_3) - \ldots - P(A_{n-1}A_n) + P(A_1A_2A_3)+\ldots+(-1)^{n+1}P(A_1A_2\ldots A_n)$

Доказательство:

1) $\Omega=A+\bar{A} \Rightarrow P(\Omega) = P(A + \bar{A}) = P(A) + P(\bar{A}) = 1$
   $P(\bar{A})=1 - P(A)$
2) $A=A+\emptyset \Rightarrow P(A)= P(A+\emptyset)=P(A)+P(\emptyset) \Rightarrow P(\emptyset) = 0$
3) $A \subset B \Rightarrow B = A + (B|A) \Rightarrow P(B) = P(A + B|A) = P(A) + P(B|A) \geq P(A)$
4) $\forall A \subset \Omega \; 0 \leq P(A) \leq P(\Omega) \leq 1$
5) $\begin{cases} A + B = A + (B|A) \\ B = (B|A) + AB \end{cases} \Rightarrow \begin{cases} P(A + B) = P(A) + P(B|A) \\ P(B) = P(B|A) + P(AB) \end{cases} \Rightarrow P(A + B) = P(A) + P(B) - P(AB)$
6) На примере трех
   P(A+B+C)=P(A)+P(B+C) – P(A|B+C)=P(A) + P(B) + P(C) – P(BC) – P(AB + AC) = P(A) + P(B) + P(C) – P(BC) – P(AB) – P(AC) + P(ABC)

16. Сформулировать и доказать теорему о свойствах условной вероятности.

Теорема:

1) $P(A|B) \geq 0$
2) $P(\Omega|B)=1$
3) $P(A_1+\ldots+A_n+\ldots|B)=P(A_1|B)+\ldots+P(A_n|B)+\ldots$

Доказательство:

1) $P(A|B) \geq 0$, так как $P(AB) \geq 0$ и $P(B) \geq 0$
2) $P(\Omega|B)=\frac{P(\Omega B)}{P(B)} = \frac{P(B)}{P(B)} = 1$
3) Пусть $A_i \cap A_j = \emptyset\ \forall i \neq j,\ i,j \geq 1$

   Тогда $(A_1 + \cdots + A_n + \cdots)B = A_1B + \cdots + A_nB + \cdots \Rightarrow$

$P(A_1+\ldots+A_n+\ldots|B)=\frac{P((A_1+\cdots+A_n+\cdots)B)}{P(B)} = \frac{P(A_1B+\cdots+A_nB+\cdots)}{P(B)} = P(A_1|B) + \cdots P(A_n|B) + \cdots$

Случайные величины

1. Что называют скалярной случайной величиной? Дайте определение функции распределения (вероятностей) скалярной случайной величины. Сформулируйте и докажите ее свойства.

Пусть (Ω,B,P) – вероятностное пространство. Скалярную функцию $\mathrm{X}(\omega)$, заданную на Ω, называют случайной величиной, если $\forall\, \mathrm{x} \in \mathbb{R}$ множество $\{\omega\colon \mathrm{X}(\omega) < \mathrm{x}\}$ есть множество элементарных исходов, для которых $\mathrm{X}(\omega) <$x является событием

Функцией распределения скалярной случайной величины X называют функцию F(x), значение которой в точке x равно вероятности события $\{\mathrm{X} < \mathrm{x}\}$, то есть события, состоящего из тех элементарных исходов ω, для которых $\mathrm{X}(\omega) < \mathrm{x} : F(x) = P\{\mathrm{X} < \mathrm{x}\}$

Свойства:

1) $0 \leq F(x) \leq 1$
2) $F(x_1) \leq F(x_2)\ \forall\ x_1 < x_2$
3) $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$

$F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1$

4) $P\{x_1 \le X \le x_1\} = F(x_1) - F(x_2)$
5) $F(x) = F(x-0),$ где $F(x-0) = \lim\limits_{y\to x-0} F(y)$

Доказательство:

1) Так как $F(x) = P\{X < x\}$, то из свойства $0 \le P(A) \le 1$ следует, что $0 \le F(x) \le 1$
2) Пусть $x_1 < x_2$. Тогда$\{X \le x_1\} \subset \{X < x_2\}$. Следовательно, по свойству вероятности 3 $P\{X < x_1\} \le P\{X < x_2\} \Rightarrow F(x_1) \le F(x_2)$
3) Поскольку $\{x < +\infty\}$ достоверное событие, то $P\{x < +\infty\} = 1$
   Так как для любой возрастающей последовательности $x_1, \dots, x_n, \dots$ $\{X < +\infty\} = \bigcup_n \{X < x_n\}$, то в силу аксиомы непрерывности $\lim\limits_{x_n\to+\infty} P\{X < x_n\} = P\{X < +\infty\} = F(+\infty) = \lim\limits_{x\to+\infty} F(x) = 1$
   Аналогично $F(-\infty)$
4) Если $x_1 < x_2$, то событие $\{X < x_2\} \cup \{x_1 \le X \le x_2\}$
   Так как события несовместны, то $P\{X < x_2\} = P\{X < x_1\} + P\{x_1 \le X \le x_2\} \Rightarrow P\{x_1 \le X \le x_2\} = P\{X < x_2\} - P\{X < x_1\} = F(x_2) - F(x_1)$
5) Пусть $x_1, \dots, x_n, \dots$ - возрастающая последовательность такая, что $\lim\limits_{n\to+\infty} x_n = X$. Событие $\{X < x\} = \bigcup_n \{X < x_n\}$. В силу аксиомы непрерывности $\lim\limits_{n\to+\infty} P\{X < x_n\} = P\{X < x\}$. То есть $\lim\limits_{x_k\to x-0} F(x_k) = F(x)$

2. Что называют скалярной случайной величиной? Дайте определение дискретной скалярной случайной величины и ее ряда распределения. Сформулируйте и докажите теорему о виде функции распределения (вероятностей) дискретной скалярной случайной величины.

Пусть (Ω,B,P) – вероятностное пространство. Скалярную функцию X(ω), заданную на Ω, называют случайной величиной, если $\forall\, \mathrm{x} \in \mathbb{R}$ множество $\{\omega\colon \mathrm{X}(\omega) < \mathrm{x}\}$ есть множество элементарных исходов, для которых $\mathrm{X}(\omega) <$x является событием

Случайная величина называется дискретной, если множество ее возможных решений конечно или счетно

Рядом распределения дискретной случайной величины X называют таблицу, состоящую из двух строк: в первой строке все возможные значения случайной величины $x_1, \dots, x_n$; во второй строчке вероятности $p_i = P\{X = x_i\}$

X	x_1	x_2	...	x_i	...	x_n
P	p_1	p_2	...	p_i	...	p_n

Теорема:

Если X – дискретная скалярная величина со множеством возможных значений $\{x_k\}_{k=1}^{N \le \infty}$, то ее функция распределения F(x) имеет вид

$$F(x) = \begin{cases} 0, & x \le x_1 \\ \sum p_k, & x_1 < x \le x_N \\ 1, & x > x_N \end{cases}$$

Доказательство:

Пусть X – дискретная случайная величина, заданная рядом распределения, причем значения $x_1, \dots, x_n$ расположены в порядке возрастания. Тогда $\forall\, X \le x_1\, \{X < x\} -$ невозможное событие и, следовательно, $F(x) = P\{X < x\} = P\{\emptyset\} = 0$

Если $x_1 < X \le x_n$, то существует k < N такой, что $x_k < X \le x_{k+1}$. Тогда $\{X < x\} = \{X = x_1\} \cup \{X = x_2\} \cup \dots \cup \{X = x_k\}$. Поскольку

$\{X < x_2\} = \{X < x_1\} \cup \{x_1 \le X < x_2\} = \emptyset \cup \{X = x_1\}$

…

$\{X < x_k\} = \{X < x_{k-1}\} \cup \{x_{k-1} \le X < x_k\} = \{X < x_{k-1}\} \cup \{X = x_{k-1}\}$,
то P$\{X < x\}$ =P$\{X = x_1\}$ + P$\{X = x_2\}$ +…+ P$\{X = x_k\}$= $p_1 + p_2 + \cdots + p_k$

Если $x > x_N$, то $\{X < x\} = \Omega$ - достоверное событие и $P\{X < x\} = 1 = F(x)$

3. Какая логическая проблема возникает при введении понятия непрерывной скалярной случайной величины и какое решение этой проблемы можно предложить? Дайте определение непрерывной скалярной случайной величин. Сформулируйте и докажите основные свойства плотности распределения вероятностей скалярной случайной величины.

Если скалярная случайная величина X может принимать любое значение из интервала $I \in \mathbb{R}$, то эту случайную величину естественно называть непрерывной. При этом возникает логическая проблема при определении Р$\{X = x \in I\}$:

Если $(P\{X = x \in I\}) > 0$, то $P\{x \complement I\} = \sum_{x \complement I} P\{X = x\} = \infty$

Если $(P\{X = x \in I\}) = 0$, то $P\{x \complement I\} = \sum_{x \complement I} P\{X = x\} = 0$

Данный парадокс объясняется некорректностью суммирования по несчетному множеству точек $x \complement I$

Непрерывной называют случайную величину X, функцию распределения которой можно представить в виде сходящегося несобственного интеграла $F(x) = \int_{-\infty}^{x} f(y)dy$. Функцию $f(x)$ называют плотностью распределения случайной величины X

Свойства:

1) $f(x) \geq 0$
2) $P\{x_1 \leq X \leq x_2\} = \int_{x_1}^{x_2} f(x)dx$
3) $\int_{-\infty}^{+\infty} f(x)dx = 1$
4) $P\{x \leq X \leq x + \Delta x\} \approx f(x)\Delta x$ в точках непрерывности плотности распределения
5) $P\{X = x\} = 0$

Доказательство:

1) Так как $F(x)$ неубывающая функция, то $f(x) = F'(x) \geq 0$
2) Так как $P\{x_1 \leq X \leq x_2\} = F(x_2) - F(x_1)$, то из определения и свойства аддитивности сходящегося несобственного интеграла
$$F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(x)dx - \int_{-\infty}^{x_1} f(x)dx = \int_{x_1}^{x_2} f(x)dx$$
3) Так как событие $\{-\infty < X < +\infty\}$ является достоверным, то $\int_{-\infty}^{+\infty} f(x)dx = F(+\infty) - F(-\infty) = 1$
4) $P\{x \leq X \leq x + \Delta x\} = F(x + \Delta x) - F(x) = \Delta F(x)$.
Если Δx мало, то $\Delta F(x) \approx dF(x) = F'(x)\Delta x = f(x)\Delta x$
5) Так как функция $F(x) = \int_{-\infty}^{x} f(y)dy$ – несобственный интеграл от плотности распределения, то $F(x)$ – непрерывная функция распределения и $\lim\limits_{\Delta x \to 0} \Delta F(x) = 0$
Тогда $P\{X = x\} = \lim\limits_{\Delta x \to 0}\big(F(x + \Delta x) - F(x)\big) = 0$

4. Что называют скалярной случайной величиной? Какая случайная величина называется дискретной? Дайте определение скалярной случайной величины, имеющей биномиальный закон распределения. Какой вид имеет ее функция распределения вероятностей?

Пусть (Ω,B,P) – вероятностное пространство. Скалярную функцию X(ω), заданную на Ω, называют случайной величиной, если $\forall\, x \in \mathbb{R}$ множество $\{\omega\colon X(\omega) < x\}$ есть множество элементарных исходов, для которых X(ω) <x является событием

Случайная величина называется дискретной, если множество ее возможных решений конечно или счетно

Дискретная случайная величина X распределена по биномиальному закону, если она принимает значения 0,1,2,…,n с вероятностью $C_n^k p^k q^{n-k}$, k=0,…,n , то есть ее закон распределения имеет вид $P\{X = k\} = P_n(k) = C_n^k p^k q^{n-k}$; k=0,…,n; q=1-p

5. Что называют скалярной случайной величиной? Какая случайная величина называется дискретной? Дайте определение скалярной случайной величины, имеющей распределение Пуассона. Какой вид имеет ее функция распределения вероятностей?

Пусть (Ω,B,P) – вероятностное пространство. Скалярную функцию X(ω), заданную на Ω, называют случайной величиной, если $\forall\, x \in \mathbb{R}$ множество $\{\omega\colon X(\omega) < x\}$ есть множество элементарных исходов, для которых X(ω) <x является событием

Случайная величина называется дискретной, если множество ее возможных решений конечно или счетно

Дискретная случайная величина X распределена по закону Пуассона, если она принимает неотрицательные целые значения с вероятностями

$\frac{\lambda^k}{k!}e^{-\lambda}; k = 0,1, \dots$, то есть ее закон распределения имеет вид $P\{X = k\} = P(k;\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}; k = 0,1, \dots$

$\lambda > 0 -$ параметр распределения пуассона

6. Дайте определение непрерывной скалярной случайной величин. Дайте определение скалярной случайной величины, имеющей равномерный закон распределения. Какой вид имеет ее функция распределения?

Пусть (Ω,B,P) – вероятностное пространство. Скалярную функцию X(ω), заданную на Ω, называют случайной величиной, если $\forall$ x $\in$ $\mathbb{R}$ множество $\{\omega\text{: X}(\omega) < \text{x}\}$ есть множество элементарных исходов, для которых X(ω) <x является событием

Если на $[a; b] \in \mathbb{R}$ множество возможных значений непрерывной случайной величины X и все они равновероятны, то говорят, что эта случайная величина распределена на $[a; b]$ равномерно.

Функция распределения имеет вид

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 1, & x > b \end{cases}$$

7. Дайте определение непрерывной скалярной случайной величины. Дайте определение нормальной случайной величины. Что называют функцией Лапласа?

Непрерывной называют случайную величину X, функцию распределения которой можно представить в виде сходящегося несобственного интеграла $F(x) = \int_{-\infty}^{x} f(y)dy$. Функцию $f(x)$ называют плотностью распределения случайной величины X

Говорят, что непрерывная случайная величина распределена по нормальному закону с параметрами m и σ^2, если ее функция распределения имеет вид

$$f(x) = \varphi_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}, -\infty < m < \infty; \sigma > 0$$

Функцией Лапласа называют интеграл $\Phi_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx$

Свойства:

1) $\Phi(-\infty) = -0{,}5 \qquad \Phi(+\infty) = 0{,}5$
2) $\Phi(0) = 0$
3) $\Phi(-x) = -\Phi(x)$
4) $\Phi(x) = P\{0 \le X \le x\}$, где $X \sim N(0{,}1)$

8. Дайте определение непрерывной скалярной случайной величин. Дайте определение скалярной случайной величины, имеющей экспоненциальный закон распределения. Какой вид имеет ее функция распределения?

Непрерывной называют случайную величину X, функцию распределения которой можно представить в виде сходящегося несобственного интеграла $F(x) = \int_{-\infty}^{x} f(y)dy$. Функцию $f(x)$ называют плотностью распределения случайной величины X

Случайная величина распределена по экспоненциальному закону с параметром $\lambda > 0$, если ее плотность распределения имеет вид

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \ge 0 \end{cases}$$

Функция распределения имеет вид

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

9. Что понимают под n-мерным случайным вектором и его функцией распределения (вероятностей)? Сформулируйте и докажите основные свойства функции распределения (вероятностей) n-мерного случайного вектора.

Совокупность случайных величин $X_1 = x_1(\omega), \dots, X_n = x_n(\omega)$, заданных на одном и том же вероятностном пространстве (Ω,B,P) называют n-мерной случайной величиной или n-мерным случайным вектором. При этом случайные величины $X_1, X_2, \dots, X_n$- координаты случайного вектора.

Функцией распределения $F(x_1, \dots, x_n) = F_{X_1,\dots,X_n}(x_1, \dots, x_n)$ случайного вектора $(X_1, \dots, X_n)$ называют функцию, значение которой в точке $(x_1, \dots, x_n) \in \mathbb{R}^n$ равно вероятности совместного осуществления событий$\{X_1 < x_1\}, \dots, \{X_n < x_n\}$, то есть

$$F(x_1, \dots, x_n) = F_{X_1,\dots,X_n}(x_1, \dots, x_n) = P\{X_1 < x_1, \dots, X_n < x_n\}$$

Свойства:

1) $0 \leq F(x_1, x_2) \leq 1$
2) $F(x_1, x_2)$ – неубывающая функция по каждому из аргументов
3) $F(-\infty, x_2) = F(x_1, -\infty) = 0$
4) $F(+\infty, +\infty) = 1$
5) $P\{a_1 \leq x_1 < b_1, a_2 \leq x_2 < b_2\} = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$
6) $F(x_1, x_2)$ – непрерывная слева в любой точке $(x_1, x_2) \in \mathbb{R}^2$ по каждому из аргументов функции
7) $F_{X_1,X_2}(x_1, +\infty) = F_{X_1}(x)$
   $F_{X_1,X_2}(+\infty, x_2) = F_{X_2}(x)$

Доказательство:

1) Аналогично одномерному
   Так как $F(x_1, \dots, x_n) = P\{X_1 < x_1, \dots, X_n < x_n\}$, то из свойства $0 \leq P(A) \leq 1$ следует, что $0 \leq F(x_1, \dots, x_n) \leq 1$
2) Аналогично одномерному
   Пусть $x_1 < x_3$ и $x_2 < x_4$.
   Тогда$\{X_1 < x_1, \dots, X_2 < x_2\} \subset \{X_3 < x_3, \dots, X_4 < x_4\}$.
   Следовательно, по свойству вероятности 3
   $P\{X_1 < x_1, \dots, X_2 < x_2\} \leq P\{X_3 < x_3, \dots, X_4 < x_4\} \Rightarrow F(x_1, x_2) \leq F(x_3, x_4)$
3) События $\{X_1 < -\infty\}$ и $\{X_2 < -\infty\}$ являются невозможными, следовательно, $F(-\infty, x_2) = P\{X_1 < -\infty, X_2 < x_2\} = 0$
   $F(x_1, -\infty) = P\{X_1 < x_1, X_2 < -\infty\} = 0$
4) Так как $\{X_1 < +\infty\}$ и $\{X_2 < +\infty\}$ достоверные события, то $\{X_1 < +\infty\} \cap \{X_2 < +\infty\}$ достоверное событие и, следовательно, $F(+\infty, +\infty) = P\{X_1 < +\infty, X_2 < +\infty\} = 1$
5) $P\{X_1 < a_1, a_2 \leq X_2 < b_2\} = P\big\{\{X_1 < a_1, X_2 < b_2\} \backslash \{X_1 < a_1, X_2 < a_2\}\big\} = F(a_1, b_2) - F(a_1, a_2)$

$$P\{X_1 < b_1, a_2 \le X_2 < b_2\} = P\big\{\{X_1 < b_1, X_2 < b_2\} \setminus \{X_1 < b_1, X_2 < a_2\}\big\}$$
$$= F(b_1, b_2) - F(b_1, a_2)$$

Тогда $P\{a_1 \le x_1 < b_1, a_2 \le x_2 < b_2\} = P\big\{\{X_1 < b_1, a_2 \le X_2 < b_2\} \setminus \{X_1 < a_1, a_2 \le X_2 < b_2\}\big\}$=$F(b_1, b_2) - F(b_1, a_2) - \big(F(a_1, b_2) - F(a_1, a_2)\big) = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$

6) Аналогично одномерному:
   Пусть $x_1, \dots, x_n, \dots$ - возрастающая последовательность такая, что $\lim\limits_{n \to +\infty} x_n = X$. Событие $\{X < x\} = \cup \{X < x_n\}$. В силу аксиомы непрерывности $\lim\limits_{n \to +\infty} P\{X < x_n\} = P\{X < x\}$. То есть $\lim\limits_{x_k \to x - 0} F(x_k) = F(x)$
7) Так как событие $\{X_2 < +\infty\}$ является достоверным событием, то $\{X_1 < x_1\} \cap \{X_2 < +\infty\} = \{X_1 < x_1\} \Rightarrow F(x_1, +\infty) = P\{X_1 < x_1\} = F_{X_1}(x)$
   Аналогично $F(+\infty, x_2) = F_{X_2}(x)$
   $F_{X_1}(x)$ и $F_{X_2}(x)$ – одномерные функции распределения случайных величин X_1 и X_2

10. Дайте определение непрерывного случайного вектора. Сформулируйте и докажите основные свойства плотности распределения вероятностей непрерывного случайного вектора.

Непрерывной двумерной случайной величиной (X,Y) называют такую двумерную случайную величину (X,Y), совместную функцию распределения которой $F(x, y) = P\{X < x, Y < y\}$ можно представить в виде сходящегося несобственного интеграла

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(t_1, t_2) dt_1 dt_2$$

Функцию $f(x, y) = F_{X,Y}(x, y)$ называют совм6стной плотностью распределения случайных величин X и Y или плотностью распределения случайного вектора (X,Y)

Свойства:

1) $f(x, y) \ge 0$
2) $P\{a_1 \le X < b_1, a_2 \le Y < b_2\} = \int_{a_1}^{b_1} dx \int_{a_2}^{b_2} f(x, y) dy$
3) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$

4) $P\{x < X < x + \Delta x, y < Y < y + \Delta y\} \approx f(x,y)\Delta x\Delta y$
5) $P\{X = x, Y = y\} = 0$
6) $P\{(x,y) \in D\} = \iint_D f(x,y)dx, dy$
7) $f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$$

Доказательство:

1) Аналогично одномерному:
Так как $F(x)$ неубывающая функция, то $f(x) = F'(x) \geq 0$
2) Аналогично одномерному:
Так как $P\{x_1 \leq X \leq x_2\} = F(x_2) - F(x_1)$, то из определения и свойства аддитивности сходящегося несобственного интеграла

$$F(x_2) - F(x_1) = \int_{-\infty}^{x_2} f(x)dx - \int_{-\infty}^{x_1} f(x)dx = \int_{x_1}^{x_2} f(x)dx$$

3) Аналогично одномерному:
Так как событие $\{-\infty < X < +\infty\}$ является достоверным, то $\int_{-\infty}^{+\infty} f(x)dx = F(+\infty) - F(-\infty) = 1$
4) Аналогично одномерному:

$P\{x \leq X \leq x + \Delta x\} = F(x + \Delta x) - F(x) = \Delta F(x)$.

Если Δx мало, то $\Delta F(x) \approx dF(x) = F'(x)\Delta x = f(x)\Delta x$

5) Аналогично одномерному:

Так как функция $F(x) = \int_{-\infty}^{x} f(y)dy$ – несобственный интеграл от плотности распределения, то $F(x)$ – непрерывная функция распределения и $\lim\limits_{\Delta x \to 0} \Delta F(x) = 0$

Тогда $P\{X = x\} = \lim\limits_{\Delta x \to 0}\big(F(x + \Delta x) - F(x)\big) = 0$

6) Обобщение (2)
Пусть $\{D_k\}_{k=1}^{N}$- разбиение, то есть $D = \bigcup_k D_k$ и
$D_i \cap D_j = \emptyset \; \forall \; i = j$
Аппроксимируем D_k элементарным прямоугольником $\Delta x_i \Delta y_i = \widetilde{D_k}$. Внутри каждого прямоугольника выберем точку $\xi_{i,j}(x_i, y_j)$ и вычислим $f\big(\xi_{i,j}\big) = f(x_i, y_j)$. Составим соответствующую сумму:$\sum_{i,j} f(x_i, y_j)\Delta x_i \Delta y_j$. При $\Delta x_i \Delta y_j \to 0$. Это будет интегральной суммой для $\iint_D f(x,y)dxdy$

С другой стороны, по свойству 4 $\sum_{i,j} f(x_i, y_j)\Delta x_i \Delta y_j = \sum_k P\{(x,y) \in \widetilde{D_k}\}$

Следовательно, $P\{(x,y) \in D\} = \iint_D f(x,y)dxdy$

7) Из свойства (7) функции распределения:

$$F_X(x) = F_{X,Y}(x,+\infty) = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{X,Y}(t_1,t_2)dt_1dt_2$$

Дифференцируя интеграл по переменному верхнему пределу и учитывая, что $f(x) = F'(x)$, имеем:

$$f_X(x) = F'(x) = F'_{X,Y}(x,+\infty) = \left(\int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{X,Y}(t_1,t_2)dt_1d\,t_2\right)'_x = \int_{-\infty}^{+\infty} f_{X,Y}(t_1,t_2)dt_2$$

Для $f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(t_1,t_2)dt_1$ доказывается аналогично

11. Дайте определение независимых случайных величин. Сформулируйте теоремы о необходимом и достаточном условии независимости для дискретных и для непрерывных случайных величин (для непрерывных с доказательством). Дайте определение независимых в совокупности случайных величин.

Случайные величины X и Y называют независимыми, если

$F_{X,Y}(x,y) = F_X(x)F_Y(y)$. В противном случае они зависимые

Теорема (для дискретных):

Дискретные случайные величины X и Y являются независимыми тогда и только тогда, когда для любых возможных значений x_i, y_j:

$P_{ij} = P\{X = x_i, Y = y_j,\} = P\{X = x_i\}P\{Y = y_j\} = P_{x_i}P_{y_j}$

Теорема (для непрерывных):

Для того, чтобы случайные величины X и Y были независимы необходимо и достаточно, чтобы для любых x, y $f_{X,Y}(x,y) = f_X(x)f_Y(y)$

Доказательство:

Пусть X и Y независимы, тогда $F_{X,Y}(x,y) = F_X(x)F_Y(y)$

Следовательно, $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \frac{dF_X(x)}{dx}\,\frac{dF_Y(y)}{dy} = f_X(x)f_Y(y)$

Пусть теперь $f_{X,Y}(x,y) = f_x(x)f_Y(y)$

$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)dudv = \int_{-\infty}^{x} f_X(u)du \int_{-\infty}^{y} f_Y(v)dv = F_X(x)F_Y(y)$

Случайные величины $X_1, \dots, X_n$, заданные на одном и том же вероятностном пространстве, называют независимыми в совокупности, если $F_{X_1,\dots X_n} = F_{X_1}(x_1) \dots F_{X_n}(x_n)$

12. Дайте определение функции от скалярной случайной величины. Выведите формулы для нахождения функции распределения (вероятностей) и плотности распределения вероятностей непрерывной случайной величины Y = φ(X) при известной функции плотности $f_X(x)$.

Случайную величину Y, которая каждому элементарному исходу ставит в соответствие число $Y(\omega) = Y(X(\omega))$ называют функцией $Y(X)$ от случайной величины

Функция распределения:

По определению $F_Y(y) = P\{Y < y\} = P\{Y(X(\omega)) < y\}$

Событие $\{Y(X(\omega)) < y\}$ эквивалентно событию $\bigcup_k \{X(\omega) \in \Delta k\}$,

где Δk – непересекающиеся промежутки из $\mathbb{R}$

Тогда по расширенной аксиоме сложения $F_Y(y) = P\{Y(X(\omega)) < y\} = \sum_k P\{X(\omega_k) \in \Delta_k\} = \sum_k \int_{\Delta_k} f_X(x)dx = \int_{\Delta} f_X(x)dx$, где $\Delta = \bigcup_k \Delta_k$

Так как $\bigcup_k \Delta_k$ определено как множество тех значений $x(\omega)$, для которых $Y(X(\omega)) < y$, то получаем $F_Y(y) = \int_{Y(X)<y} f_X(x)dx$

Плотность распределения:

Пусть функция $\varphi(x) = y = Y(x)$ непрерывна и монотонна. Тогда существует обратная функция $x = \psi(y) = \varphi^{-1}(y)$

Тогда событие $\{\varphi(X(\omega) < y)\}$ эквивалентно событию $\{X(\omega) < \psi(y)\}$ для возрастающей функции($\{X(\omega) > \psi(y)\}$ для убывающей функции)

Тогда

$$P\{Y < y\} = P\{\varphi(X(\omega)) < y\} = \begin{cases} P\{X(\omega) < \psi(y)\}, & \varphi(x) \text{ возрастает} \\ P\{X(\omega) > \psi(y)\}, & \varphi(x) \text{ убывает} \end{cases} = \begin{cases} F_X(\psi(y)), & \varphi(x) \text{ возрастает} \\ 1 - F_X(\psi(y)), & \varphi(x) \text{ убывает} \end{cases}$$

Функция плотности может быть найдена как производная от функции распределения:

$$f_Y(y) = F'_Y(y) = \begin{cases} f_X\big(\psi(y)\big)\psi'(y), \varphi(x) \text{ возрастает} \\ -f_X\big(\psi(y)\big)\psi'(y), \varphi(x) \text{ убывает} \end{cases}$$

Оба эти случая можно записать в виде $f_Y(y) = f_X(\psi(y))|\psi'(y)|$

13. Дайте определение функции от двумерной случайной величины. Запишите формулу для нахождения функции распределения (вероятностей) случайной величины Z = ψ(X, Y) при известной совместной плотности распределения вероятностей f_{XY} (x, y) случайного вектора (X, Y). Выведите формулу свертки.

Случайную величину $Y = \varphi(X_1, X_2) = \varphi(X_1(\omega), X_2(\omega))$ называют функцией от двумерной случайной величины (X_1, X_2)

Если (X_1, X_2) двумерная непрерывная случайная величина с плотностью распределения $f_{X,Y}(x, y)$, то функцию распределения случайной величины $Y = \varphi(X_1, X_2)$ можно найти по формуле

$F_Y(y) = \iint_{\varphi(X_1,X_2)<y} f_{X_1,X_2}(X_1, X_2)dX_1dX_2$, где область интегрирования состоит из всех значений (X_1, X_2) для которых $\varphi(X_1, X_2) < y$

Формула свертки:

Пусть X_1 и X_2 – независимые случайные величины и $Y = X_1 + X_2$, то есть $\varphi(x_1, x_2) = x_1 + x_2$. Тогда $f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$ и

$$F_Y(y) = \iint_{x_1+x_2<y} f_{X_1,X_2}(x_1, x_2)dx_1dx_2 = \iint_{x_1+x_2<y} f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2$$
$$= \int_{-\infty}^{+\infty} f_{X_1}(x_1)dx_1 \int_{-\infty}^{y-x_1} f_{X_2}(x_2)dx_2 = \int_{-\infty}^{+\infty} F_{X_2}(y - x_1)f_{X_1}(x_1)dx_1$$

Тогда дифференцируя по y под знаком интеграла имеем:

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X_2}(y - x)f_{X_1}(x)dx = f_{X_2} \cdot f_{X_1}$$

14. Дайте определение математического ожидания скалярной случайной величины и приведите его содержательную интерпретацию.

Сформулируйте и докажите основные свойства математического ожидания.

Математическим ожиданием MX дискретной случайной величины X называют сумму произведений значений x_i случайной величины и вероятностей $p_i = P\{X = x_i\}$, с которыми случайная величина принимает эти значения:

$$MX = \sum_i x_i p_i$$

Пусть на прямой расположена система материальных точек с массами p_i ($\sum_i p_i = 1$) и пусть x_i- координата i-ой точки. Тогда центр масс системы будет иметь координату $\bar{X} = \frac{\sum p_i x_i}{\sum p_i} = \frac{\sum p_i x_i}{1} = \sum p_i x_i$, совпадающую с математическим ожиданием MX случайной величины X

Математическим ожиданием MX непрерывной случайной величины X называют интеграл $MX = \int_{-\infty}^{+\infty} x f(x) dx$

Так же, как и в дискретном случае, математическое ожидание непрерывной случайной величины можно интерпретировать как центр масс стержня, плотность массы которого в точке x равна $f(x)$

Свойства:

1) Если случайная величина X принимает всего одно значение C с вероятностью 1, то MC=C
2) M$(aX + b) = aMX + b$, где a,b константы
3) $M(X_1 + X_2) = MX_1 + MX_2$
4) $M(X_1 X_2) = MX_1 MX_2$, если случайные величины X_1 и X_2 независимы

Доказательство:

1) Если случайная величина X принимает всего одно значение C с вероятностью 1, то по определению MC=C$\cdot 1 = C$
2) Найдем математическое ожидание случайной величины $Y = aX + b$ ($Y = Y(x) = ax + b$):

$$MY = M(aX+b) = \int_{-\infty}^{+\infty}(ax+b)f_X(x)dx = a\int_{-\infty}^{+\infty} x f_X(x)dx + b\int_{-\infty}^{+\infty} f_X(x)dx = aMX + b\cdot 1 = aMX + b$$

3) Пусть теперь $Y = X_1 + X_2 (Y = Y(x_1, x_2) = x_1 + x_2)$
   Тогда

$$MY = M(X_1+X_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1+x_2)f_{x_1x_2}(x_1,x_2)dx_1dx_2 = \int_{-\infty}^{+\infty} x_1\left(\int_{-\infty}^{+\infty} f_{x_1x_2}(x_1,x_2)dx_2\right)dx_1$$
$$+\int_{-\infty}^{+\infty} x_2\left(\int_{-\infty}^{+\infty} f_{x_1x_2}(x_1,x_2)dx_1\right)dx_2 = \int_{-\infty}^{+\infty} x_1 f_{x_1}(x_1)dx_1 + \int_{-\infty}^{+\infty} x_2 f_{x_2}(x_2)dx_2 = MX_1 + MX_2$$

4) Так как X_1 и X_2 независимые случайные величины, то воспользовавшись формулой для математического ожидания функции случайной величины $(Y = X_1X_2)$ имеем:

$$MY = M(X_1X_2) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1x_2\, f_{X_1,X_2}(x_1,x_2)dx_1dx_2$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_1x_2\, f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2$$
$$= \left(\int_{-\infty}^{+\infty} x_1\, f_{X_1}(x_1)dx_1\right)\left(\int_{-\infty}^{+\infty} x_2\, f_{X_2}(x_2)dx_2\right) = MX_1MX_2$$

15. Дайте определение дисперсии скалярной случайной величины. Сформулируйте и докажите основные свойства дисперсии.

Дисперсией DX случайной величины X называют математическое ожидание квадрата отклонения случайной величины X от ее среднего значения, то есть $DX = M(X - MX)^2$

Свойства:

1) Если случайная величина X принимает всего одно значение C с вероятностью 1, то DC=0
2) $D(aX+b) = a^2DX$
3) $DX = MX^2 - (MX)^2$
4) $D(X+Y) = DX + DY$, если X и Y независимые
5) $D(X+Y) = DX + DY + 2cov(X,Y)$ для любых X и Y

Доказательство:

1) $DX = M(X-C)^2 = (C-C)^2\cdot 1 = 0$
2) $DY = M(Y-MY)^2 = M(aX+b-M(aX+b))^2 = M(aX+b-aMX-b)^2 = M(a(X-MX))^2 = M(a^2(X-MX)^2) = a^2M(X-MX)^2 = a^2DX$
3) $DX = M(X-MX)^2 = M(X^2 - 2XMX + (MX)^2) = MX^2 - 2(MX)^2 + (MX)^2 = MX^2 - (MX)^2$
4) Пусть $\dot{X} = X - MX \qquad \dot{Y} = Y - MY$

$D(X+Y) = M(X+Y-M(X+Y))^2 = M((X-MX)+(Y-MY))^2 = M(X-MX)^2 + 2M\big((X-MX)(Y-MY)\big) + M(Y-MY)^2 = DX + 2\big(M\dot{X}M\dot{Y}\big) + DY = DX + DY$, поскольку $M\dot{X} = 0$ и $M\dot{Y} = 0$

5) $D(X+Y) = M((X+Y)-M(X+Y))^2 = M(X-MX)^2 + 2M\big((X-MX)(Y-MY)\big) + M(Y-MY)^2 = DX + DY + 2cov(X,Y)$

16. Дайте определение ковариации двух скалярных случайных величин. Сформулируйте и докажите основные свойства ковариации.

Ковариацией $cov(X_1, X_2)$ случайных величин X_1 и X_2 называют математическое ожидание произведения случайных величин

$$\dot{X}_1 = X_1 - MX_1 \quad \text{и} \quad \dot{X}_2 = X_2 - MX_2$$

$$cov(X_1, X_2) = M\big(\dot{X}_1, \dot{X}_2\big) = M((X_1 - MX_1)(X_2 - MX_2))$$

Свойства:

1) $cov(X, X) = DX$
2) $cov(X_1, X_2) = 0$, если X_1 и X_2 независимые
3) Если $Y_1 = a_1X_1 + b_1$ и $Y_2 = a_2X_2 + b_2$, то $cov(Y_1, Y_2) = a_1a_2cov(X_1, X_2)$
4) $-\sqrt{DX_1DX_2} \le cov(X_1, X_2) \le \sqrt{DX_1DX_2}$
5) $|cov(X_1, X_2)| = \sqrt{DX_1DX_2}$ тогда и только тогда, когда $X_2 = aX_1 + b$
6) $cov(X_1, X_2) = M(X_1X_2) - MX_1MX_2$

Доказательство:

1) По определению ковариации и дисперсии $cov(X, X) = M(X - MX)^2 = DX$
2) Если случайные величины X_1 и X_2 независимы, то $cov(X_1, X_2) = M\big((X_1 - MX_1)(X_2 - MX_2)\big) = \big(M(X_1 - MX_1)(X_2 - MX_2)\big) = 0$
3) Пусть $Y_1 = a_1X_1 + b_1$ и $Y_1 = a_1X_1 + b_1$
   Тогда $cov(Y_1, Y_2) = M\big((Y_1 - MY_1)(Y_2 - MY_2)\big) = M\big((a_1X_1 + b_1 - a_1MX_1 - b_1)(a_2X_2 + b_2 - a_2MX_2 - b_2)\big) = M\big(a_1a_2(X_1 - MX_1)(X_2 - MX_2)\big) = a_1a_2cov(X_1, X_2)$
4) Рассмотрим дисперсию случайной величины $Y_x = xX_1 - X_2$, x – произвольное число.
   В силу свойств дисперсии и свойства 3 ковариации
   $DY_x = D(xX_1) + 2cov(xX_1, X_2) + D(-X_2) = x^2DX_1 - 2xcov(X_1, X_2) + DX_2$
   Так как $DY_x \ge 0$, то дискриминант
   D= $(-2cov(X_1, X_2))^2 - 4DX_1DX_2 \le 0$, следовательно

$$|cov(X_1, X_2)| \leq \sqrt{DX_1DX_2}$$

5) Пусть выполнено равенство $|cov(X_1, X_2)| = \sqrt{DX_1DX_2}$

   Значит дискриминант D= $\left(2cov(X_1, X_2)\right)^2 - 4DX_1DX_2$ равен нулю, и уравнение $DY_x = 0$ имеет решение, которое обозначим a

   Тогда случайная величина $Y_a = aX_1 - X_2$ принимает всего одно значение, допустим b, поэтому $X_2 = aX_1 - b$

   Пусть теперь выполнено $X_2 = aX_1 - b$

   Тогда, в соответствии со свойством 1 дисперсии $DY_a = 0$, а значит, дискриминант неотрицателен, следовательно, он равен нулю, а значит $|cov(X_1, X_2)| = \sqrt{DX_1DX_2}$

6) $cov(X_1, X_2) = M\big((X_1 - MX_1)(X_2 - MX_2)\big) = M(X_1X_2 - X_2MX_1 - X_1MX_2 + MX_1MX_2) = M(X_1X_2) - M(X_2MX_1) - M(X_1MX_2) + MX_1MX_2 = M(X_1X_2) - MX_1MX_2$

17. Дайте определение коэффициента корреляции двух скалярных случайных величин. Сформулируйте основные свойства коэффициента корреляции.

Коэффициентом корреляции случайных величин X и Y называют число $\rho = \rho(X, Y)$, определяемое равенством $\rho = \frac{cov(X,Y)}{\sqrt{DXDY}}$

Свойства:

1) $\rho(X, X) = 1$
2) Если случайные величины X и Y независимы, то $\rho(X, Y) = 0$
3) $\rho\left(a_1X_1 + b_{1,}a_2X_2 + b_2\right) = \pm\rho(X_1, X_2)$ (+ если знаки a одинаковы)
4) $-1 \leq \rho(X, Y) \leq 1$
5) $|\rho(X, Y)| = 1$ тогда и только тогда, когда X и Y связаны линейной связью

18. Дайте определения сходимости по вероятности. Сформулируйте и докажите теоремы о 1-м и 2-м неравенствах Чебышева.

Если последовательность $X_1, X_2, \dots, X_n, \dots$ случайных величин для любого $\varepsilon > 0$ удовлетворяет условию $\lim_{n\to\infty} P\{|x_n| < \varepsilon\} = 1$, то говорят о сходимости этой последовательности к нулю по вероятности

Теорема:

Для каждой неотрицательной случайной величины X, имеющей математическое ожидание MX, при любом $\varepsilon > 0$ справедливо первое неравенство Чебышева $P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon}$

Доказательство:

Рассмотрим непрерывную случайную величину X с плотностью распределения $f(x)$. Поскольку случайная величина X является неотрицательной, то $MX = \int_0^{+\infty} xf(x)dx$.

Так как подынтегральное выражение неотрицательно, то при уменьшении области интегрировании интеграл может только уменьшится, то есть

$$MX = \int_0^{\varepsilon} xf(x)dx + \int_{\varepsilon}^{+\infty} xf(x)dx \geq \int_{\varepsilon}^{+\infty} xf(x)dx$$

Заменяя в подынтегральном выражении сомножитель x на ε имеем

$$\int_{\varepsilon}^{+\infty} xf(x)dx \geq \varepsilon \int_{\varepsilon}^{+\infty} f(x)dx$$

Последний интеграл представляет собой вероятность события $x \geq \varepsilon$, поэтому $MX \geq \varepsilon P\{X \geq \varepsilon\} \Rightarrow P\{X \geq \varepsilon\} \leq \frac{MX}{\varepsilon}$

Аналогично и для дискретной случайной величины

Теорема:

Для каждой случайной величины X, имеющей дисперсию $DX = \sigma^2$, при любом $\varepsilon > 0$ справедливо второе неравенство Чебышева

$$P\{|X - MX| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

Доказательство:

Рассмотрим случайную величину $Y = (X - MX)^2$. Поскольку

$Y \geq 0$ и $MY = (X - MX^2) = DX = \sigma^2$, то мы можем применить к ней первое неравенство Чебышева, в котором ε заменено на ε^2

$$P\{|X - MX| \geq \varepsilon\} = P\{(X - MX)^2 \geq \varepsilon^2\} = P\{Y \geq \varepsilon^2\} \leq \frac{MY}{\varepsilon^2} = \frac{DX}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

19. Дайте определения сходимости по вероятности. Дайте определения закона больших чисел (слабого). Сформулируйте и докажите закон больших чисел в форме Чебышева.

Если последовательность $X_1, X_2, \dots, X_n, \dots$ случайных величин для любого $\varepsilon > 0$ удовлетворяет условию $\lim\limits_{n\to\infty} P\{|x_n| < \varepsilon\} = 1$, то говорят о сходимости этой последовательности к нулю по вероятности

Последовательность $X_1, X_2, \dots, X_n, \dots$ случайных величин удовлетворяет закону больших чисел (слабому),если для любого $\varepsilon > 0$

$$P\left\{\left|\frac{1}{n}\sum_{1}^{n} X_i - \frac{1}{n}\sum_{1}^{n} M_i\right| \geq \varepsilon\right\} \xrightarrow[n\to\infty]{} 0$$

Теорема:

Если последовательность $X_1, X_2, \dots, X_n, \dots$ независимых случайных величин такова, что существует $MX_i = m_i$ и $DX = \sigma_i^2$, причем

$\sigma_i^2 \leq \mathrm{C} < +\infty$, то для этой последовательности выполнен закон больших чисел. При этом говорят также, что к этой последовательности применен закон больших чисел в форме Чебышева.

Доказательство:

Рассмотрим последовательность случайных величин

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

В силу свойств математического ожидания и дисперсии имеем

$$MY_n = M\left(\frac{1}{n}\sum_{1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} m_i$$

$$DY_n = D\left(\frac{1}{n}\sum_{1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2{}_i \leq \frac{Cn}{n^2} = \frac{C}{n}$$

Применяя теперь второе неравенство Чебышева к случайным величинам Y_n, получаем для любого $\varepsilon > 0$

$$P\{|Y_n - MY_n| \geq \varepsilon\} = P\left\{\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n M_i\right| \geq \varepsilon\right\} \leq \frac{C}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0$$

Следовательно, для последовательности $X_1, X_2, \dots, X_n, \dots$ выполняется закон больших чисел

20. Сформулируйте центральную предельную теорему (частный случай). Как следует понимать термин «асимптотическая нормальность»? Сформулируйте интегральную теорему Муавра-Лапласа.

Пусть $X_1, X_2, \dots, X_n, \dots$ - последовательность независимых одинаково распределенных случайных величин, $MX_n = m$, $DX_n = \sigma^2$. Тогда

$$P\left\{\frac{S_n - mn}{\sqrt{n\sigma^2}} < x\right\} \xrightarrow[n\to\infty]{} \Phi(x)$$

Где $\Phi(x)$ – функция стандартного нормального распределения

Пусть $X_1, X_2, \dots, X_n, \dots$ - последовательность независимых одинаково распределенных случайных величин, $MX_n = m$

Обозначим $S_n = \sum X_i$. Тогда закон больших чисел для этой последовательности имеет вид $\frac{1}{n}(S_n - nm) \xrightarrow[n\to\infty]{} 0$

Если дополнительно известно, что существует дисперсия $DX_n = \sigma^2$, то оказывается, что случайная величина S_n распределена асимптотически нормально, то есть $S_n \xrightarrow[n\to\infty]{} N(nm, n\sigma^2)$

Теорема:

Пусть S_n – суммарное число успехов в n испытаниях по схеме Бернулли с вероятностью успеха p и вероятностью неудачи q=1-p. Тогда с ростом n последовательность функций распределения случайных величин $\frac{S_n - np}{\sqrt{npq}}$ сходится к функции стандартного нормального распределения, то есть

$$P\left\{\frac{S_n - np}{\sqrt{npq}} < \text{x}\right\} \xrightarrow[n\to\infty]{} \Phi(x)$$

21. Вычислите математическое ожидание и дисперсию случайной величины имеющих распределение: Пуассона, равномерное, экспоненциальное, нормальное.

Пуассона:

$$p = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$MX = \sum X_i p_i = \cdots = \lambda$$

$$DX = MX^2 - (MX)^2 = \cdots = \lambda$$

Равномерное:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$$

$$MX = \int_{-\infty}^{+\infty} x f(x) dx = \cdots = \frac{b-a}{2}$$

$$DX = \int_{-\infty}^{+\infty} (X - MX)^2 f(x) dx = \cdots = \frac{(b-a)^2}{12}$$

Экспоненциальное:

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}$$

$$MX = \int_{-\infty}^{+\infty} x f(x) dx = \cdots = \frac{1}{\lambda}$$

$$DX = \int_{-\infty}^{+\infty} (X - MX)^2 f(x) dx = \cdots = \frac{1}{\lambda^2}$$

Нормальное:

$$f(x) = \varphi_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

$$MX = m$$

$$DX = \sigma^2$$

22. Дайте определения сходимости по вероятности. Дайте определения закона больших чисел (слабого). Сформулируйте и докажите теорему Бернулли.

Если последовательность $X_1, X_2, \dots, X_n, \dots$ случайных величин для любого $\varepsilon > 0$ удовлетворяет условию $\lim_{n\to\infty} P\{|x_n| < \varepsilon\} = 1$, то говорят о сходимости этой последовательности к нулю по вероятности

Последовательность $X_1, X_2, \dots, X_n, \dots$ случайных величин удовлетворяет закону больших чисел (слабому),если для любого $\varepsilon > 0$

$$P\left\{\left|\frac{1}{n}\sum_1^n X_i - \frac{1}{n}\sum_1^n M_i\right| \geq \varepsilon\right\} \xrightarrow[n\to\infty]{} 0$$

Теорема:

Пусть проводится n испытаний по схеме Бернулли и Y_n – общее число успехов в n испытаниях. Тогда наблюдаемая частота успехов $r_n = \frac{Y_n}{n}$ сходится по вероятности к вероятности успеха p в одном испытании, то есть для любого $\varepsilon > 0$ $\quad P\{|r_n - p| \geq \varepsilon\} \xrightarrow[n\to\infty]{} 0$

Доказательство:

Пусть X_i – число успехов в i-ом испытании Бернулли, то есть $X_i \in \{0,1\}$. Тогда частоту успехов в n испытаниях можно определить в виде $r_n = \frac{1}{n}\sum_1^n X_i$, причем $MX_i = p$ и $DX_i = pq$. Значит выполняются все условия следствия, из которого вытекает утверждение теоремы

Математическая статистика

1. Сформулируйте основную задачу математической статистики. Какое противоречие принципиально присуще математической статистики. Дайте определение генеральной совокупности.

Основная задача математической статистики – разработка методов получения научно обоснованных выводов о массовых явлениях и процессах по данным наблюдений или экспериментов.

Если воспользоваться понятием «наиболее правдоподобное значение», то в соответствии с ним следует считать $p = \frac{m}{N}$. Но можно рассмотреть и иные «наиболее правдоподобные значения» и считать, что $p \in (p_1, p_2)$, где $0 < p_1 < \frac{m}{N} < p_2 < 1$.

При этом, чем больше величина $I \triangleq p_1 - p_2$, тем с большей уверенностью мы можем утверждать, что $p \in (p_1, p_2)$, и тем меньшую информацию относительно p мы получаем – противоречие, которое принципиально присуще математической статистике

В математической статистике множество возможных значений случайной величины X называют генеральной совокупностью X

2. Что называют выборкой, случайной выборкой, статистикой? Дайте определение выборочного начального момента k-го порядка. Как принято называть начальный момент первого порядка?

Совокупность независимых случайных величин $X_{1,\dots,}X_n$, каждая из которых имеет то же распределение, что и случайная величина X, называется случайной выборкой $\overrightarrow{X_n}$ из генеральной совокупности X.

Любое возможное значение $\overrightarrow{x_n} = (x_1, \dots, x_n)$ случайной выборки $\overrightarrow{X_n}$ называется выборкой из генеральной совокупности X

Любую функцию $g(\overrightarrow{X_n})$ от случайной выборки $\overrightarrow{X_n}$ называют статистикой

Статистику $\hat{\mu}_k(\overrightarrow{X_n}) = \frac{1}{n}\sum_1^n X_i^k$ называют выборочным начальным моментом k-го порядка

Выборочный начальный момент первого порядка называют выборочным средним и обозначают $\bar{X}$

3. Что называют выборкой, случайной выборкой, статистикой? Дайте определение выборочного центрального момента k-го порядка. Как принято называть центральный момент второго порядка?

Совокупность независимых случайных величин $X_{1,\dots,}X_n$, каждая из которых имеет то же распределение, что и случайная величина X, называется случайной выборкой $\overrightarrow{X_n}$ из генеральной совокупности X.

Любое возможное значение $\overrightarrow{x_n} = (x_1, \dots, x_n)$ случайной выборки $\overrightarrow{X_n}$ называется выборкой из генеральной совокупности X

Любую функцию $g(\overrightarrow{X_n})$ от случайной выборки $\overrightarrow{X_n}$ называют статистикой

Статистику $\hat{\upsilon}_k\left(\overrightarrow{X_n}\right) = \frac{1}{n}\sum_1^n (X_i - \bar{X})^k$ называют выборочным центральным моментом k-го порядка

Выборочный центральный момент второго порядка называют выборочной дисперсией и обозначают $\hat{\sigma}^2(\overrightarrow{X_n})$

4. Дайте определение точечной оценки. Дайте определение несмещенной и состоятельной точечных оценок.

Точечной оценкой параметра $\theta \in \Theta$ называют любую статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$, выборочное значение которой $\hat{\theta}(\overrightarrow{x_n})$ можно было бы считать приближенным значением параметра θ

Статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют несмещенной оценкой параметра θ, если ее математическое ожидание совпадает с θ, то есть $M\hat{\theta}\left(\overrightarrow{X_n}\right) = \theta$ для любого фиксированного n

Статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют состоятельной оценкой параметра θ, если с ростом объема выборки n она сходится по вероятности к оцениваемому параметру θ, то есть $\hat{\theta}\left(\overrightarrow{X_n}\right) \xrightarrow[n\to\infty]{P} \theta$

5. Дайте определение точечной оценки. Дайте определение несмещенной и эффективной точечных оценок.

Точечной оценкой параметра $\theta \in \Theta$ называют любую статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$, выборочное значение которой $\hat{\theta}(\overrightarrow{x_n})$ можно было бы считать приближенным значением параметра θ

Статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют несмещенной оценкой параметра θ, если ее математическое ожидание совпадает с θ, то есть $M\hat{\theta}\left(\overrightarrow{X_n}\right) = \theta$ для любого фиксированного n

Если в некотором классе несмещенных оценок параметра θ, имеющих конечную дисперсию, существует такая оценка $\hat{\theta}\left(\overrightarrow{X_n}\right)$, что для всех остальных оценок $\tilde{\theta}(\overrightarrow{X_n})$ из данного класса выполняется неравенство

$D\hat{\theta}\left(\overrightarrow{X_n}\right) < D\tilde{\theta}(\overrightarrow{X_n})$, то оценку $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют эффективной в данном классе

Эффективную оценку в классе всех несмещенных оценок называют эффективной оценкой

6. Дайте определение точечной оценки. Дайте определение эффективной и состоятельной точечных оценок.

Точечной оценкой параметра $\theta \in \Theta$ называют любую статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$, выборочное значение которой $\hat{\theta}(\overrightarrow{x_n})$ можно было бы считать приближенным значением параметра θ

Если в некотором классе несмещенных оценок параметра θ, имеющих конечную дисперсию, существует такая оценка $\hat{\theta}\left(\overrightarrow{X_n}\right)$, что для всех остальных оценок $\tilde{\theta}(\overrightarrow{X_n})$ из данного класса выполняется неравенство

$D\hat{\theta}\left(\overrightarrow{X_n}\right) < D\tilde{\theta}(\overrightarrow{X_n})$, то оценку $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют эффективной в данном классе

Эффективную оценку в классе всех несмещенных оценок называют эффективной оценкой

Несмещенная: мат ожидание = параметру

Статистику $\hat{\theta}\left(\overrightarrow{X_n}\right)$ называют состоятельной оценкой параметра θ, если с ростом объема выборки n она сходится по вероятности к оцениваемому параметру θ, то есть $\hat{\theta}\left(\overrightarrow{X_n}\right) \xrightarrow[n\to\infty]{P} \theta$

7. Дайте определения выборочной функции распределения, эмпирической функции распределения и эмпирической функции плотности распределения.

Функцию $\hat{F}\left(x, \overrightarrow{X_n}\right) = \frac{n(x,\overrightarrow{X_n})}{n}$ называют выборочной функцией распределения

Для каждой реализации $\overrightarrow{x_n}$ случайной выборки $\overrightarrow{X_n}$ реализацию $\hat{F}(x, \overrightarrow{x_n})$ выборочной функции распределения $\hat{F}\left(x, \overrightarrow{X_n}\right)$, определенную на выборке $\overrightarrow{x_n}$ называют эмпирической функцией распределения и обозначают $F_n(x) \triangleq \hat{F}(x, \overrightarrow{x_n}) \equiv \frac{n(x,\overrightarrow{X_n})}{n}$

Функция $f_n(x)$, которая во всех точках интервала J_i, $i = 1, \dots, m$ принимает значения $\frac{n_i}{n\Delta}$, вне интервала J равна нулю $f_n(x) = \begin{cases} \frac{n_i}{n\Delta}, & x \in J \\ 0, & x \notin J \end{cases}$ называется эмпирической плотностью распределения.

8. Дайте определения функции правдоподобия и оценки максимального правдоподобия. Что такое уравнения правдоподобия?

Функцией правдоподобия называют функцию вида

$L\left(X_1, \dots, X_2, \vec{\theta}\right) = \prod_1^n p(X_i, \theta)$, где $p(X_i, \theta) = P\{X_i = x_i\}$ для дискретных и $p(X_i, \theta) = f(X_i, \theta)$ для непрерывных случайных величин

Оценкой максимального правдоподобия параметра θ называют статистику $\hat{\vec{\theta}}\left(\overrightarrow{X_n}\right)$, значения $\hat{\vec{\theta}}$ которой для любой выборки $\overrightarrow{x_n}$ удовлетворяет условию $L\left(\overrightarrow{x_n}, \hat{\vec{\theta}}\right) = \max\limits_{\vec{\theta} \in \Theta} L\left(\overrightarrow{x_n}, \vec{\theta}\right)$

Уравнения правдоподобия: $\frac{\partial \ln L(\overrightarrow{x_n}, \vec{\theta})}{\partial \theta_k} = 0$, где $\vec{\theta} = (\theta_1, \dots, \theta_r); k = 1, \dots, r$

9. Изложите идею метода максимального правдоподобия построения точечных оценок параметров законов распределения дискретных случайных величин.

Рассмотрим функцию правдоподобия случайной выборки $\overrightarrow{X_n}$ плотность распределения которой $p(X_i, \theta)$ известна с точностью до θ

$$L(X_1, \dots, X_2, \vec{\theta}) = \prod_{1}^{n} p(X_i, \theta)$$

Если функция $L(\overrightarrow{x_n}, \hat{\vec{\theta}})$ дифференцируема как функция аргумента $\vec{\theta}$ при лбом значении $\overrightarrow{x_n}$ и максимум этой функции достигается во внутренней точке из Θ, то значение точечной оценки максимального правдоподобия удовлетворяет уравнению правдоподобия $\frac{\partial L(\overrightarrow{x_n}, \vec{\theta})}{\partial \theta_k} = 0$

Так как при логарифмировании точки экстремума остаются теми же, уравнение упрощается, то уравнение правдоподобия записывают в виде: $\frac{\partial \ln L(\overrightarrow{x_n}, \vec{\theta})}{\partial \theta_k} = 0$

10. Сформулируйте теорему Рао (неравенство Рао-Крамера).

Пусть рассматриваемая параметрическая модель является регулярной, то есть ее можно дифференцировать по параметру под знаком интеграла, зависящего от параметра, и $\hat{\vec{\theta}}(\overrightarrow{X_n})$ – несмещенная оценка неизвестного параметра θ. Тогда имеет место неравенство

$$D\, \hat{\vec{\theta}}(\overrightarrow{X_n}) \geq \frac{1}{nI(\theta)},$$

Где $I(\theta) = M(\frac{\partial \ln f(X,\theta)}{\partial \theta})^2$ – количество информации по Фишеру в одном наблюдении, а $f(t, \theta)$ – плотность распределения генеральной совокупности X в случае непрерывной статистической модели и вероятность события $\{X = t\}$ в случае дискретной статистической модели.

11. Что называют показателем эффективности по Рао-Крамеру? Какую точечную оценку называют эффективной по Рао-Крамеру?

Величину $e(\theta) = \frac{1}{nI(n) D\hat{\theta}(\overrightarrow{X_n})}$ называют показателем эффективности по Рао-Крамеру

Из $D\hat{\theta}\left(\overrightarrow{X_n}\right) \geq \frac{1}{nI(\theta)}$ следует, что для любой несмещенной оценки $\hat{\theta}\left(\overrightarrow{X_n}\right)$ параметра θ выполняется неравенство $0 \leq e(\theta) \leq 1$

Несмещенную оценку $\hat{\theta}\left(\overrightarrow{X_n}\right)$ параметра θ называют эффективной по Рао-Крамеру, если ее показатель эффективности $e(\theta) = 1$

12. Дайте определение интервальной оценки. Как определяется вероятность совершения ошибки при построении γ-доверительного интервала. Дайте определение центральной статистики.

Интервальная оценка – оценка, позволяющая получить вероятностную характеристику точности оценивания неизвестного значения параметра θ известного закона распределения $F_X(x, \theta)$ наблюдаемой случайной величины X по данным случайной выборки $\overrightarrow{X_n}$

Вероятность α совершения ошибки при нахождении интервальной оценки параметра θ с коэффициентом доверия γ равна $\alpha = 1 - \gamma$, то есть

$$\alpha = P\left\{\theta \notin \underline{\theta}(\overrightarrow{X_n}) < \theta < \overline{\theta}(\overrightarrow{X_n})\right\}$$

Статистику $g(\overrightarrow{X_n}, \theta)$ называют центральной, если ее закон распределения не зависит от θ, то есть функция распределения вероятностей $F_g(t) = P\left\{g(\overrightarrow{X_n}, \theta) < t\right\}$ не зависит от θ

13. Сформулируйте задачу построения интервальной оценки. Какие допущения используются при ее построении?

Допущения:

1) Функция распределения $F_g(t)$ центральной статистики $g(\overrightarrow{X_n}, \theta)$ является непрерывной и возрастающей
2) Заданы положительные числа α_1 и α_2 такие, что $\alpha = \alpha_1 + \alpha_2$, то есть коэффициент доверия $\gamma = 1 - \alpha_1 - \alpha_2$
3) Для любой выборки $\overrightarrow{x_n}$ из генеральной совокупности X функция $g(\overrightarrow{x_n}, \theta)$ является непрерывной и возрастающей(убывающей) функцией по θ

Построение:

Из (1): Для любого $q \in (0,1)$ существует единственный корень h_q уравнения $F_g(t) = q$, который называют квантилью уровня q функции распределения случайной величины

Из (2): $1 - \alpha_1 - \alpha_2 = \gamma = F_g(h_{1-\alpha_2}) - F_g(h_{\alpha_1}) = P\{h_{\alpha_1} < g(\overrightarrow{X_n}, \theta) < h_{1-\alpha_2}\}$ для любого θ

Из (3): для каждой реализации случайной выборки каждое из уравнений $g(\overrightarrow{x_n}, \theta) = h_{\alpha_1}$ и $g(\overrightarrow{x_n}, \theta) = h_{1-\alpha_2}$ имеют единственные решения $\underline{\theta}(\overrightarrow{x_n})$ и $\overline{\theta}(\overrightarrow{x_n})$

При этом $h_{\alpha_1} < g(\overrightarrow{x_n}, \theta) < h_{1-\alpha_2} \Leftrightarrow \underline{\theta}(\overrightarrow{x_n}) < \theta < \overline{\theta}(\overrightarrow{x_n})$

Таким образом, $\gamma = 1 - \alpha_1 - \alpha_2 = P\{h_{\alpha_1} < g(\overrightarrow{X_n}, \theta) < h_{1-\alpha_2}\} = P\{\underline{\theta}(\overrightarrow{X_n}) < \theta < \overline{\theta}(\overrightarrow{X_n})\}$ и $(\underline{\theta}(\overrightarrow{X_n})$ и $\overline{\theta}(\overrightarrow{X_n}))$ – искомая интервальная оценка

14. Сформулируйте задачу построения интервальной оценки. Пусть X ~ N(m,σ²). Укажите вид доверительного интервала для математического ожидания при известной дисперсии.

Рассмотрим статистику $g(\vec{X}_n, m) = \frac{m-\overline{X}}{\sigma}\sqrt{n}$

$$u_{\alpha_1} = g(\vec{x}_n, \underline{m}(\vec{x}_n)) = \frac{\sqrt{n}}{\sigma}(\underline{m}(\vec{x}_n) - \overline{X}) \Rightarrow \underline{m}(\vec{x}_n) = \overline{X} + \frac{\sigma}{\sqrt{n}} u_{\alpha_1}$$

$$u_{1-\alpha_2} = g(\vec{x}_n, \overline{m}(\vec{x}_n)) = \frac{\sqrt{n}}{\sigma}(\overline{m}(\vec{x}_n) - \overline{X}) \Rightarrow \overline{m}(\vec{x}_n) = \overline{X} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha_2}$$

15. Сформулируйте задачу построения интервальной оценки. Пусть X ~ N(m,σ²). Укажите вид доверительного интервала для математического ожидания при неизвестной дисперсии.

Рассмотрим статистику $g(\vec{X}_n, m) = \frac{m-\overline{X}}{S(\vec{X}_n)}\sqrt{n}$

Данная статистика является центральной и распределена по закону Стьюдента с n-1 степенью свободы

$$t_{\alpha_1}(n-1) \leq \frac{m - \overline{X}}{S(\vec{X}_n)}\sqrt{n} \leq t_{1-\alpha_2}(n-1)$$

$$\underline{m}(\vec{X}_n) = \overline{X} - \frac{S(\vec{X}_n)}{\sqrt{n}} t_{1-\alpha_1}(n-1)$$

$$\overline{m}(\vec{X}_n) = \overline{X} + \frac{S(\vec{X}_n)}{\sqrt{n}} t_{1-\alpha_2}(n-1)$$

16. Сформулируйте задачу построения интервальной оценки. Пусть X ~ N(m, σ^2). Укажите вид доверительного интервала для дисперсии.

Рассмотрим статистику $g(\vec{X}_n, \sigma) = \frac{(n-1)S^2(\vec{X}_n)}{\sigma^2}$

Данная статистика имеет χ^2- распределение с n-1 степенью свободы и является центральной

Приведенная статистика является убывающей функцией параметра σ. Поэтому $\underline{\sigma}(\vec{X}_n) = \frac{S(\overrightarrow{X_n})\sqrt{n-1}}{\sqrt{\chi^2_{1-\alpha_1}(n-1)}}$ $\overline{\sigma}(\vec{X}_n) = \frac{S(\overrightarrow{X_n})\sqrt{n-1}}{\sqrt{\chi^2_{\alpha_1}(n-1)}}$

17. Дайте определение статистической и параметрической гипотез. Дайте определение простой и сложной статистических гипотез

Статистической гипотезой называют любое утверждение о функции распределения случайной величины X

Параметрической гипотезой называют любое утверждение о параметре θ функции распределения $F(t, \theta)$ случайной величины X. При этом если θ – скаляр, то речь идет об однопараметрических гипотезах, а если вектор – то о многопараметрических гипотезах

Статистическую гипотезу H называют простой, если она имеет вид $H: \vec{\theta} = \vec{\theta}_0$, где $\vec{\theta}_0$ – некоторое заданное значение параметра

Статистическую гипотезу H называют сложной, если она имеет вид $H: \vec{\theta} \in D$, где D – некоторое множество значений параметра θ, состоящее более, чем из одного элемента

18. Что называют критерием проверки статистической гипотезы и как его задают? Что представляет собой решающее правило?

Критерием проверки гипотез называют правило, по которому по данным выборки $\overrightarrow{x_n}$ принимается решение о справедливости либо основной, либо конкурирующей гипотезы

Критерий задают с помощью критического множества $W \in \mathbb{R}^n$, являющегося подмножеством выборочного пространства χ_n случайной выборки $\overrightarrow{X_n}$

Решение принимают следующим образом:

- Если выборка $\overrightarrow{x_n}$ принадлежит критическому множеству W, то отвергают основную гипотезу H_0 и принимают альтнрнативную гипотезу H_1
- Если выборка $\overrightarrow{x_n}$ принадлежит критическому множеству W, то отвергают альтернативную гипотезу H_1 и принимают основную гипотезу H_0

19. Какие ошибки возможны при проверке статистической гипотезы? Что понимают под мощностью и уровнем значимости критерия?

Ошибки:

- Принять гипотезу H_1, когда верна H_0 – ошибка первого рода. Вероятность ее совершения $\alpha = P\{\vec{X}_n \in W|H_0\}$
- Принять гипотезу H_0, когда верна H_1 – ошибка второго рода. Вероятность ее совершения $\beta = P\left\{\vec{X}_n \in \overline{W} = \frac{\chi}{W}|H_1\right\}$

Вероятность совершения ошибки первого рода α называют уровнем значимости критерия

Величину $1 - \beta$, равную вероятности отвергнуть основную гипотезу H_0, когда она неверна, называют мощностью критерия

20. Простые параметрические гипотезы: постановка задачи. Какую функцию называют отношением правдоподобия и как определяется критическое множество?

Пусть $\vec{X}_n$ – случайная выборка объема n из генеральной совокупности непрерывной случайной величины X, плотность распределения которой $f(t, \theta)$ зависит от неизвестного параметра θ. Рассмотрим две простые гипотезы H_0: $\theta = \theta_0$ и H_1: $\theta = \theta_1$

Для решения поставленной задачи проверки двух простых параметрических гипотез введем следующую функцию случайной выборки $\vec{X}_n$:

$$\varphi(\vec{X}_n) = \frac{L(\vec{X}_n, \theta_1)}{L(\vec{X}_n, \theta_0)}$$

$$L(\vec{X}_n\theta) = \prod f(X_i, \theta)$$

Статистика $\varphi(\vec{X}_n)$ представляет собой отношение функций правдоподобия при истинности альтернативной и основной гипотез соответственно. Ее называют отношением правдоподобия

Критическое множество W определяется следующим соотношением:

$$W = \{\vec{x}_n \in \chi: \varphi(\vec{x}_n) \geq C_\varphi\}$$

Где константу C_φ выбирают из условия $P\{\varphi(\vec{x}_n) \geq C_\varphi | H_0\} = \alpha$, которое обеспечивает заданное значение уровня значимости α и, для непрерывной модели, может быть записано в виде

$$\iint\limits_{\varphi(t_1,\dots,t_n) \geq C_\varphi} L(t_1, \dots, t_n, \theta_0) dt_1 \dots dt_n = \alpha$$

21. Что понимают под размером и функцией мощности критерия при проверке сложных параметрических гипотез?

Ошибки:

$\alpha(\theta) = P\{(X_1, \dots, X_n) \in W | \theta\},\ \theta \in \Theta_0$ – первого рода

$\beta(\theta) = P\{(X_1, \dots, X_n) \in \overline{W} | \theta\},\ \theta \in \Theta_1$ – второго рода

Максимально возможное значение вероятности совершения ошибки первого рода $\alpha = \max\limits_{\theta \in \Theta_0} \alpha(\theta)$ называют размером критерия

Функцию $M(\theta) = P\{(X_1, \dots, X_n) \in W | \theta\}$, определяющую значение вероятности отклонения основной гипотезы H_0 в зависимости от истинного значения параметра θ, называют функцией мощности критерия

22. Дайте определение равномерно наиболее мощного критерия. Каким основным свойством обладает этот критерий? Почему?

Если существует критерий, который при данном фиксированном размере α максимизирует функцию мощности $M(\theta) = P\{(X_1, \dots, X_n) \in W|\theta\}$ по всем возможным критериям одновременно при всех θ из множества Θ_1, то такой критерий называют равномерно наиболее мощным

Вероятности совершения ошибок первого и второго рода связаны с функцией мощности следующими соотношениями:

$$\alpha(\theta) = M(\theta), \ \ \theta \in \Theta_0$$

$$\beta(\theta) = 1 - M(\theta), \ \ \theta \in \Theta_1$$

Тем самым равномерно наиболее мощный критерий, если он существует, минимизирует вероятность совершения ошибки второго рода $\beta(\theta)$ (при фиксированном размере α) одновременно при всех $\theta \in \Theta_1$