1 Предварительные сведения

Опр. 1. *Пусть $\Omega = \{\omega\}$ - произвольное множество, а $\mathcal{F}$ - σ-**алгебра** его подмножеств, то есть система множетсв, таких что:*

1. $\Omega \in \mathcal{F}$
2. *Если* $A \in \mathcal{F}$, *то* $\bar{A} := \Omega - A \in \mathcal{F}$
3. *Если* $A_1, A_2, \ldots \in \mathcal{F}$, *то* $\bigcup_i A_i \in \mathcal{F}$ *и* $\bigcap_i A_i \in \mathcal{F}$

Пример. *Система всех подмножеств $\mathcal{F}$ - σ-алгебра*

Пример. $\{0, \Omega\}$ - *σ-алгебра*

Опр. 2. *Пусть $\Omega = \mathbb{R}$, а $\mathcal{F}$ - наименьшая сигма-алгебра, содержащая все интервалы (α, β). Такая $\mathcal{F}$ обозначается $\mathfrak{B}(\mathbb{R})$ и называется **борелевской сигма-алгеброй**.*

Опр. 3. *Мера μ, определенная на $\mathcal{F}$, называется **сигма-аддитивной**, если это неотрицательная функция, $\mu(A) \geq 0$ для $A \in \mathcal{F}$, и она удовлетворяет условию сигма-аддитивности, то есть:*

$$\mu(\bigcup_i A_i) = \sum_i \mu(A_i), \ A_i \in \mathcal{F}, A_i \cap A_j \underset{i \neq j}{=} \varnothing$$

Опр. 4. *Мера μ называется **сигма-конечной**, если $\exists$ множетсва $A_i \in \mathcal{F}$ такие, что $\bigcup_i A_i = \Omega$ и $\mu(A_i) < \infty$*

Пример (Считающая мера)**.** *Пусть Ω - счетное, $\mathcal{F}$ - множество всех подмножеств Ω. Положим для $A \in \mathcal{F}$*

$$\mu(A) := \{\text{числу точек } \Omega, \text{ попавших в } A\}$$

Такая мера называется считающей, она сигма-конечна.

Пример (Лебегова мера)**.** *Пусть $\Omega = \mathbb{R}, \mathcal{F} = \mathfrak{B}(\mathbb{R})$. $\exists!$ мера μ на $\mathfrak{B}(\mathbb{R})$ такая, что*

$$\mu((\alpha, \beta]) = \beta - \alpha$$

Это мера Лебега, она сигма-конечна.

Опр. 5. $(\Omega, \mathcal{F})$ - ***измеримое пространство***. $(\Omega, \mathcal{F}, \mu)$ - ***пространство с мерой***.

Опр. 6. *Если $\mu(\Omega) = 1$, то μ - **вероятностная мера**, она обозначается через* P.

Опр. 7. *Тройка $(\Omega, \mathcal{F}, \mathrm{P})$ - **вероятностное пространство**.*

Опр. 8. *Измеримая функция $\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}.\mathfrak{B}(\mathbb{R}))$ (то есть $\forall B \in \mathfrak{B}(\mathbb{R})$ $\xi^{-1}(B) := (\omega : \xi(\omega) \in B) \in \mathcal{F}$) называется **случайной величной**.*

*Измеримая функция $\phi : (\mathbb{R}.\mathfrak{B}(\mathbb{R})) \to (\mathbb{R}.\mathfrak{B}(\mathbb{R}))$ называется **борелевской**.*

Опр. 9. *Рассмотрим сл. в. $\xi \in \mathbb{R}^1$. Для $x \in \mathbb{R}^1$ функция $F(x) = \mathrm{P}(\omega : \xi(\omega) \leq x) = \mathrm{P}(\xi \leq x)$ называется **функцией распределения**.*

Опр. 10. *Мера $\mathrm{P}_\xi(A) := \mathrm{P}(\omega : \xi(\omega) \in A)$, $A \in \mathfrak{B}(\mathbb{R})$, называется **распределением** случайной величины ξ.*

Тогда $F(x) = \mathrm{P}_\xi((-\infty, x])$, то есть $\underline{\mathrm{P}_\xi \text{ определяет } F(x)}$.

Обратно: $\mathrm{P}(\alpha < \xi \leq \beta) = F(\beta) - F(\alpha)$, и $\exists!$ вероятнаостная мера P_ξ такая, что $\mathrm{P}_\xi((\alpha, \beta]) = F(\beta) - F(\alpha)$, то есть $\underline{F(x) \text{ определяет } \mathrm{P}_\xi}$.

Опр. 11. *Пусть на $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ задана σ-конечная мера μ. Если $\exists$ борелевская функция $f(x), f(x) \geq 0$ такая, что:*

$$\mathrm{P}_\xi(A) = \int_A f(x)\mu(dx) \ \forall A \in \mathfrak{B}(\mathbb{R})$$

*то $f(x)$ называется **плотностью вероятности случайной величины по мере** μ.*

Если μ - мера Лебега, то $f(x)$ - обычная плотность вероятности сл. в. ξ, введенная на 2-ом курсе. Если же ξ дискретна со значениями $x_1, x_2, \ldots$, а μ - считающая мера, сосредоточенная в этих точках, то, очевидно,

$$\mathrm{P}_\xi(A) = \int_A \mathrm{P}(\xi = x)\mu(dx) \ \forall A \in \mathfrak{B}(\mathbb{R})$$

Последнее равенство означает, что у дискретной случайной величины ξ есть плотность вероятности $f(x) = \mathrm{P}(\xi = x), \ x = x_1, x_2, \ldots$ по считающей мере. (При $x \neq x_1, x_2, \ldots$ значения не важны, их можно положить равными 0)

Опр. 12. ***Математическим ожиданием** случайной величины ξ называется число*

$$\mathrm{E}\xi = \int_\Omega \xi(\omega)\mathrm{P}(d\omega)$$

(в предположении, что $\int_\Omega |\xi(\omega)|\mathrm{P}(d\omega) < \infty$, иначе говорим, что мат. ожидание $\nexists$)

Если $f(x)$ - плотность вероятности случайной величины ξ по мере μ, а $\phi(x)$ - борелевская функция, то

$$\mathrm{E}\phi(\xi) = \int_\mathbb{R} \phi(x)\mathrm{P}_\xi(dx) = \int_\mathbb{R} \phi(x)f(x)\mu(dx)$$

В частности, если ξ - абсолютно непрерывная случайная величина в терминологии 2-го курса (то есть μ - мера Лебега), то пишем

$$\mathrm{E}\phi(\xi) = \int_\mathbb{R} \phi(x)f(x)dx$$

Разумеется, только в случае $\int_\mathbb{R} |\phi(x)|f(x)dx < \infty$. Если же ξ дискретна со значениями $x_1, x_2, \ldots$ и соответствующими вероятностями, то

$$\mathrm{E}\phi(\xi) = \sum_{i \geq 1} \phi(x_i)p_i \text{ (если ряд сходится абсолютно)}$$

Опр. 13. *Обозначим $\mathfrak{B}(\mathbb{R}^K)$ борелевскую σ-алгебру подмножеств $\mathbb{R}^K$. Вектор $\xi = (\xi_1, \ldots, \xi_k)^T$ называется **k-мерным случайным вектором**, если ξ - измеримое отображение $\xi : (\Omega, \mathcal{F}) \to (\mathbb{R}^K, \mathfrak{B}(\mathbb{R}^K))$*

Известно: ξ - случайный вектор $\Leftrightarrow$ каждая компонента ξ_i - одномерная случайная величина.

Опр. 14. ***Функция распределения случайного вектора** ξ:*

$$F(x_1, \ldots, x_K) = \mathrm{P}(\xi_1 \leq x_1, \ldots, \xi_K \leq x_K), x_i \in \mathbb{R}$$

Опр. 15. ***Распределение:*** $\mathrm{P}_\xi(A) = \mathrm{P}(\omega : \xi(\omega) \in A), \ A \in \mathfrak{B}(\mathbb{R}^K)$.

Опр. 16. ***Плотность вероятности вектора*** ξ ***по мере*** μ *(μ определена на элементах $\mathfrak{B}(\mathbb{R}^K)$) - борелевская функция $f(x) \geq 0, x = (x_1, \ldots, x_K)$ такая, что:*

$$\mathrm{P}_\xi(A) = \int_A f(x)\mu(dx),\ \forall A \in \mathfrak{B}(\mathbb{R}^K)$$

Опр. 17. *Случайные величины $\{\xi_1, \ldots, \xi_K\}$* ***независимы****, если*

$$P(\xi_1 \in A_1, \ldots, \xi_K \in A_K) = \prod_{i=1}^{K} \mathrm{P}(\xi_i \in A_i)\ \forall A_i \in \mathfrak{B}(\mathbb{R})$$

Бесконечная последовательность будет последовательностью независимых величин, если каждая конечная подпоследовательность независима.

Необходимые и достаточные условия независимости.
Рассмотрим $x = (x_1, \ldots, x_K) \in \mathbb{R}^K$

1. $F(x) = F_{\xi_1}(x_1)F_{\xi_2}(x_2)\ldots F_{\xi_K}(x_K)\ \forall x \in \mathbb{R}^K$
2. *Если $\exists$ плотность $f(x)$: $f(x) = f_{\xi_1}(x_1)f_{\xi_2}(x_2)\ldots f_{\xi_K}(x_K)$ для μ-почти всех $x \in \mathbb{R}^K$*

В заключении Раздела 1 поговорим о сходимости случайных векторов.

Пусть случайные векторы $\xi, \xi_1, \xi_2, \ldots$ размера K со значениями в $(\mathbb{R}, \mathfrak{B}(\mathbb{R}^K))$ определены на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathrm{P})$. Пусть $|\cdot|$ означает Евклидову норму вектора, то есть $|\xi| = \sqrt{\sum_{i=1}^K \xi_i^2}$.

Опр. 18. *Говорят, что последовательность ξ_n сходится* ***слабо*** *к ξ, (пишем $\xi_n \xrightarrow{w} \xi,\ n \to \infty$) если для любой непрерывной и ограниченной $g : \mathbb{R}^K \to \mathbb{R}^1$*

$$\int_{\mathbb{R}^K} g(x)\mathrm{P}_n(dx) \to \int_{\mathbb{R}^K} g(x)\mathrm{P}(dx),\ n \to \infty \tag{1}$$

Здесь P_n и P - распределения соотвественно ξ_n и ξ.

В вероятностных терминах: $\xi_n \xrightarrow{w} \xi,\ n \to \infty \Leftrightarrow$ математическое ожидание $\mathrm{E}\xi_n \to \mathrm{E}\xi$

Опр. 19. *Обозначим $F_n(x),\ F(x),\ x = (x_1, \ldots, x_n)$ как функции распределения векторов ξ_n и ξ, тогда сходимостью* ***в основном*** *называют*

$$F_n(x) \Rightarrow F(x),\ \text{то есть } F_n(x) \to F(x)\ \forall x \in C(F) \tag{2}$$

Пусть $\phi_n(t)$ и $\phi(t), t \in \mathbb{R}^K$, будут характеристические функции ξ_n и ξ, то есть $\phi(t) := \mathrm{E}e^{it^T\xi}$.

$$\phi_n(t) \to \phi(t)\ \forall t \in \mathbb{R}^K,\ n \to \infty \tag{3}$$

Опр. 20. *Если выполнено любое из соотношений (1) - (3) будем писать*

$$\xi_n \xrightarrow{d} \xi,\ n \to \infty \tag{4}$$

И говорить, что $\{\xi_n\}$ сходится к ξ ***по распределению****.*

Замечание. *Сходимость (4) не следует из сходимости* $\xi_{i_n} \xrightarrow{d} \xi_i,\ i = 1, \ldots, K$, *компонент векторов* ξ_n *и* ξ

Рассмотрим двумерный вектор $(-\xi, \xi),\ \xi \sim N(0,1)$. $(-\xi) \xrightarrow{w} \xi$, *так как одна и та же функция распределения и плотность, значит есть покомпонентная сходимость. Почему нет слабой сходимости двумерного вектора? По теореме с прошлого семестра: если* $\xi_1 \xrightarrow{w} \xi_2$, *то* $\sum_i \xi_{1_i} \xrightarrow{w} \sum_i \xi_{2_i}$, *но* $-\xi + \xi = 0 \not\xrightarrow{w} \xi + \xi = 2\xi$, *так как разные функции распределения.*

Опр. 21. *Говорят, что последовательность* $\{\xi_n\}$ *сходится по вероятности к вектору* ξ *(пишут* $\xi_n \xrightarrow{\mathrm{P}} \xi,\ n \to \infty$*), если*

$$\mathrm{P}(|\xi_n - \xi| > \epsilon) \to 0,\ n \to \infty,\ \forall \epsilon > 0 \tag{5}$$

Понятно, что сходимость (5) эквивалентна сходимости компонент $\xi_{i_n} \xrightarrow{\mathrm{P}} \xi_i,\ \forall i = 1, \ldots, K$
(5) $\Rightarrow$ (4), но обратное верно только в частных случаях, например:

$$\text{Если } \xi_n \xrightarrow{d} c = const, \text{ то } \xi_n \xrightarrow{\mathrm{P}} c$$

Опр. 22. *Говорят, что последовательность* $\{\xi_n\}$ *сходится* ***п.н. (почти наверное или с вероятностью единица)*** *и пишут* $\xi_n \xrightarrow{п.н.} \xi,\ n \to \infty$, *если*

$$\mathrm{P}(w : \xi_n(w) \to \xi(\omega)) = 1 \tag{6}$$

Задача. *Если* ξ_n *и* ξ *скаляры, и* $\xi_n \xrightarrow{п.н.} \xi$, *то верно ли, что* $\xi_n^3 \xrightarrow{п.н.} \xi^3$*?*

Да, верно, так как у них одно и то же множество сходимости.
Сходимость (6) влечет (5), а значит верно следующее:

$$\xi_n \xrightarrow{\text{п.н.}} \xi \Rightarrow \xi_n \xrightarrow{\mathrm{P}} \xi \Rightarrow \xi_n \xrightarrow{d} \xi,\ n \to \infty$$

Теорема 1 (Теорема непрерывности)**.**
Пусть векторы $\{\xi_n\},\ \xi$ *определены на* $(\Omega, \mathcal{F}, \mathrm{P}),\ \xi_n, \xi \in \mathbb{R}^K$. *Пусть* $A \in \mathfrak{B}(\mathbb{R}^K)$, *и* $\mathrm{P}(\xi \in A) = 1$ *(то есть* A *- носитель* ξ*). Пусть борелевская* $H : \mathbb{R}^K \to \mathbb{R}^1$, *и* $H(x)$ *непрерывна на множестве* A. *Тогда:*

1. *Если* $\xi_n \xrightarrow{d} \xi$, *то* $H(\xi_n) \xrightarrow{d} H(\xi),\ n \to \infty$
2. *Если* $\xi_n \xrightarrow{\mathrm{P}} \xi$, *то* $H(\xi_n) \xrightarrow{\mathrm{P}} H(\xi),\ n \to \infty$
3. *Если* $\xi_n \xrightarrow{п.н.} \xi$, *то* $H(\xi_n) \xrightarrow{п.н.} H(\xi),\ n \to \infty$

Доказательство. Докажем пункт 3. Остальные будут доказаны на практических занятиях.
Итак, в силу непреревности функции $H(x)$ на A:

$$(\omega : \xi_n(\omega) \to \xi(\omega)) \cap (\omega : \xi(\omega) \in A) \subseteq (\omega : H(\xi_n(\omega)) \to H(\xi(\omega)))$$

Значит,

$$1 \underset{\xi_n \xrightarrow{\text{п.н.}} \xi}{=} \mathrm{P}(\xi_n(\omega) \to \xi(\omega)) \underset{A-\text{носитель}}{=} \mathrm{P}(\xi_n(\omega) \to \xi(\omega), \xi(\omega) \in A) \le \mathrm{P}(H(\xi_n(\omega)) \to H(\xi(\omega)))$$

$\square$

Пусть на $(\Omega, \mathcal{F}, \mathrm{P})$ задана бесконечная последовательность случайных величин $\xi_1, \xi_2, \ldots$.

1. Если $\{\xi_i\}$ независимы и одинаково распределены (н.о.р.) с конеченым средним, $\mathrm{E}|\xi_1| < \infty$, то

$$n^{-1}\sum_{i=1}^{n}\xi_i \xrightarrow{\text{п.н.}} \mathrm{E}\xi_1,\ n\to\infty \tag{7}$$

   Соотношение (7) - усиленный закон больших чисел Колмогорова

2. Если $\{\xi_i\}$ некоррелированные сл.в., может быть, разнораспределенные, но с одинаковым средним $m = \mathrm{E}\xi_i,\ \mathrm{D}\xi_i \le c < \infty$, то

$$n^{-1}\sum_{i=1}^{n}\xi_i \xrightarrow{\mathrm{P}} m = \mathrm{E}\xi_i,\ n\to\infty \tag{8}$$

   Соотношение (8) - слабый закон больших чисел.

3. Если $\{\xi_i\}$ - н.о.р. сл.в., $\mathrm{E}\xi_1 = m,\ 0 < \mathrm{D}\xi_1 = \sigma^2 < \infty$, то

$$\frac{1}{\sqrt{n}\sigma}\sum_{i=1}^{n}(\xi_i - m) \xrightarrow{d} \xi \sim N(0,1),\ n\to\infty \tag{9}$$

   Соотношение (9) - центральная предельная теорема (точнее, ее вариант). То есть

$$n^{1/2}(\overline{\xi} - m) \xrightarrow{d} N(0,\sigma^2), \text{ где } \overline{\xi} := n^{-1}\sum_{i=1}^{n}\xi_i$$

2 Примеры статистических задач. Статистическая модель.

Пример (Оценка среднего).

Для многих изделий одна из основных характеристик - срок службы. Срок службы обычно случаен и заранее неизвестен. Опыт показывает: для однородного процесса производства сроки службы $x_1, x_2, \dots, x_n$ соответственно 1-го, 2-го, и т.д. изделий можно рассматривать как реализации н.о.р. сл.в. $X_1, X_2, \dots, X_n$

(Будем предполагать, что на некотором вероятностном пространстве $(\Omega, \mathcal{F}, \mathrm{P})$ определена бесконечная последовательность $X_1, X_2, \dots$ и $X_1, \dots, X_n$ - её первые n членов)

Интересующий нас параметр, определяющий (в какой-то мере) срок службы отождествим с $\theta = \mathrm{E}X_1$.

Одна из стандарстных статистических задач состоит в том, чтобы выяснить, чему равно θ. Вот возможное решение.

В силу Усиленного Закона Больших Чисел (УЗБЧ) Колмогорова

$$\overline{X} := \frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{\text{п.н.}} \mathrm{E}X_1 = \theta,\ n\to\infty$$

Возьмем n готовых изделий и проверим их. Пусть $x_1, x_2, \dots, x_n$ сроки службы готовых изделий. Это реализации сл.в. $X_1, \dots, X_n$. Естественно ожидать, что $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$ при больших n оокажется близким к θ.

<u>Это задача точечного оценивания параметра</u>

$X_1, \ldots, X_n$ - случайные наблюдения; $\overline{X}$ - (борел. ф-ция) статистическая оценка, это случайная величина $\overline{x}$ - реализация оценки, с ней работают на практике.

Ясно, что нужны оценки, которые в среднем близки θ. Тогда и реализации будут близки.

Пусть в частности,

$$\mathrm{P}(X_1 \leq t) = \begin{cases} 0, \ t \leq 0 \\ 1 - e^{-t/\theta}, \ t > 0 \end{cases} \qquad \text{параметр } \theta > 0$$

То есть $X_1 \sim Exp(\frac{1}{\theta})$, и $\mathrm{E}_\theta X_1 = \theta$.

Тогда $\overline{X}$ оптимальна при любом конечном $n \geq 1$ в следующем смысле:

1. $$\mathrm{E}_\theta \overline{X} = \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}_\theta X_i = \theta \ \forall \theta > 0$$

   Это свойство несмещенности качественно: реализации $\overline{X}$ группируются вокруг θ

2. $$\mathrm{D}_\theta \overline{X} \leq \mathrm{D}_\theta \hat{\theta}_n \ \forall \theta > 0 \text{ и любой несмещенной оценки } \hat{\theta}_n = \hat{\theta}_n(x_1, \ldots, x_n)$$

   Это свойство также качественно: реализации $\overline{X}$ в среднем лежат ближе к θ, чем у других $\hat{\theta}_n$

Пример (Проверка однородности данных)**.**

Некоторый эксперимент проводится сначала m раз в условиях A, а затем n раз в условиях B. Пусть $x_1, \ldots, x_m$ - результаты в условиях A, $y_1, \ldots, y_n$ - в условиях B.

Например, влияет ли некоторый препарат на развитие растений, лекарство на анализы больного и т.п.

Будем считать $\{x_i\}$ реализациями н.о.р. сл.в., $\{X_i\}$ с ф.р. $X_1 \sim F_X(x) = \mathrm{P}(X_1 \leq X)$. Пусть $\{y_j\}$ - реализации н.о.р. сл.в. $\{Y_j\}$, ф.р. $Y_1 \sim F_Y(x)$.

Последовательности $\{X_i\}$ и $\{Y_j\}$ независимы.

Интерпретируем поставленную задачу как проверку гипотезы

$$H : F_X = F_Y$$

. Предположение о том, что условия B дают иной результат интерпретируем как гипотезу (альтернативную к H)

$$K : F_X \neq F_Y$$

<u>Важно</u>: ни F_X, ни F_Y неизвестны!

Оценкой $F_X(x)$ возьмем

$$\hat{F}_{mX}(x) = \frac{1}{m} \sum_{i=1}^{m} I(X_i \leq x), \ x \in \mathbb{R}$$

Это "хорошая"оценка, т.к. в силу УЗБЧ

$$\hat{F}_{mX}(x) = \frac{1}{m} \sum_{i=1}^{m} I(X_i \leq x) \xrightarrow{\text{п.н.}} \mathrm{E} I(X_1 \leq x) = F_X(x)$$

(У нас $\{X_i\}$ и $\{Y_j\}$ определены на одном $(\Omega, \mathcal{F}, \mathrm{P})$)

Теорема Гливенко-Кантелли.

$$\sup_x |\hat{F}_{mX}(x) - F_X(x)| \xrightarrow{п.н.} 0,\ m \to \infty$$

Очевидно, если гипотеза H верна, то величина

$$\mathrm{D}_{mn} := \sup_x |\hat{F}_{mX} - \hat{F}_{nY}| \text{ мала при больших } m,\ n$$

Получаем естественное правило:

Если $\mathrm{D}_{mn} \leq c$, то H принять;

Если $\mathrm{D}_{mn} > c$, то H опровергнуть и принять K.

Но как выбрать константу c?