БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Механико-математический факультет

Кафедра функционального анализа

ЗАДАНИЯ К ЛАБОРАТОРНЫМ РАБОТАМ ПО КУРСУ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В ДВУХ ЧАСТЯХ

Часть 2.

Для студентов специальности Н.01.01.00.
"Математика"

Минск 1999

Составители: кандидат физ.-мат. наук, доцент В.Ф.Жданович, доктор физ.-мат. наук, профессор Н.В.Лазакович, кандидат физ.-мат. наук Н.Я.Радыно, кандидат физ.-мат. наук С.П.Стащуленок

Утверждено Советом механико-математического факультета Белорусского государственного университета
21 января 1999г., протокол № 4

Учебное издание

ЖДАНОВИЧ Витольд Феликсович
ЛАЗАКОВИЧ Николай Викторович
РАДЫНО Николай Яковлевич
СТАЩУЛЕНОК Сергей Павлович

Задания к лабораторным работам
по курсу теории вероятностей и математической статистики в двух частях
Часть 2. Для студентов специальности Н.01.01.00.
"Математика".

Ответственный за выпуск Н.Я.Радыно
Подписано к печати . Формат 60 × 84/16. Бум.тип № 3.
Печать офсетная.
Усл.печ.л. . Усл.кр. . Уч.-изд.л. .
Тираж 200 экз. Заказ № . Бесплатно.

Белгосуниверситет, 220050,Минск, пр.Ф.Скорины,4.
Отпечатано на ротапринте Белгосуниверситета.
220050, Минск, Бобруйская,7.

1 Лабораторная работа №1 (2 часа)

ТЕМА: *Функции от случайных величин.*

Необходимые понятия и теоремы: борелевская функция, функция от случайной величины, σ – алгебра, порожденная случайной величиной, распределение вероятностей, плотность вероятности.

Литература: [1] стр. 55-56, 85-86, 91-93; [2] стр.146-159; [3] стр.74-84; [5] стр.60-61, 89-92; [7] стр.248-261; [11] стр. 53-55.

Задание 1.1 *Плотностью вероятности случайной величины ξ является функция $f_\xi(x) = f(x)$. Найти плотность вероятности $f_\eta(y)$, если*

1) $\eta = \xi + 1$	$(-\infty < x < +\infty)$;	*5)* $\eta = \xi^3$	$(0 < x < +\infty)$;
2) $\eta = 2\xi$	$(-a < x < a)$;	*6)* $\eta = \xi^{-1}$	$(0 < x < +\infty)$;
3) $\eta = -2\xi$	$(a < x < b)$;	*7)* $\eta = \sqrt{\xi}$	$(0 < x < +\infty)$;
4) $\eta = A\xi + B$	$(a < x < b)$;	*8)* $\eta = \exp(\xi)$	$(0 < x < +\infty)$.

Задание 1.2 *Плотность вероятности $f_\xi(x) = f(x)$. Найти плотность вероятности $f_\eta(y)$, если*

1) $\eta = \exp(-\xi)$	$(0 < x < +\infty)$;	*5)* $\eta = \operatorname{tg}\xi$	$(-\pi/2 < x < \pi/2)$;
2) $\eta = \ln\xi$	$(0 < x < +\infty)$;	*6)* $\eta = \xi^2$	$(-\infty < x < +\infty)$;
3) $\eta = \operatorname{arctg}\xi$	$(-\infty < x < +\infty)$;	*7)* $\eta = \xi^{-2}$	$(0 < x < +\infty)$;
4) $\eta = \cos\xi$	$(-\infty < x < +\infty)$;	*8)* $\eta = \lvert\xi\rvert$	$(-\infty < x < +\infty)$.

Задачи.

1. Колесу, ось которого горизонтальна, придается вращение, которое затухает вследствие трения; фиксированный радиус останавливаясь, образует с горизонтом случайный угол φ, который

равномерно распределен в пределах от 0° до 360°. Найти среднее значение и дисперсию конца радиуса a от горизонтального диаметра.

2. В каком случае вероятнее истратить более трех минут на ожидание транспорта: если ехать на автобусе (интервал движения 4 минуты) или дважды садиться в поезд метро (интервал движения 2 минуты). Считать время ожидания равномерно распределенным в интервале.

3. Пусть случайная величина ξ имеет плотность распределения $f(x)$. Найти плотность распределения случайной величины

$$\int\limits_{-\infty}^{\xi} f(x)\,dx.$$

4. Длины сторон ξ и η треугольника — независимые случайные величины с функциями распределения $F_1(x)$ и $F_2(x)$. Найти функцию распределения третьей стороны, если угол между сторонами равен постоянному числу α.

5. Пусть ξ и η независимые, нормальные $(0,\sigma^2)$ случайные величины. Доказать, что $\zeta = \xi^2 + \eta^2$, $\varkappa = \xi/\eta$ также независимы.

6. Доказать, что если величины ξ и η независимы и распределены по закону χ^2 с параметрами m и n, то величины $\zeta = \xi^2 + \eta^2$, $\varkappa = \xi/\eta$ также независимы.

7. Случайная величина ξ имеет непрерывную функцию распределения $F(x)$. В результате n независимых наблюдений над ξ получены следующие значения $x_1 < x_2 < \ldots < x_n$. Найти плотность распределения величины

$$\eta = \frac{F(x_n) - F(x_2)}{F(x_n) - F(x_1)}.$$

8. Случайные величины ξ и η независимы, их плотности распределения соответственно равны

$$f_\xi(x) = \frac{1}{\pi\sqrt{1-x^2}}, \ |x| < 1; \ f_\eta(x) = xe^{-\frac{x^2}{2}}, \ x > 0.$$

Доказать, что величина $\xi\eta$ нормально распределена.

9. Пусть ξ и η независимы и имеют плотность распределения

$$f_\xi(x) = f_\eta(x) = xe^{-x}, \ x > 0.$$

Доказать, что отношение $\zeta = \dfrac{\xi}{\xi + \eta}$ распределено равномерно на отрезке $[0, 1]$.

10. Случайные величины ξ и η независимы и равномерно распределены на $[-1, 1]$. Вычислить вероятность того, что корни уравнения

$$x^2 + \xi x + \eta = 0$$

вещественны.

11. Пусть случайная величина ξ имеет распределение Коши

$$f_\xi(x) = \frac{1}{\pi(1+x^2)}.$$

Найти $M \min\{|\xi|, 1\}$.

12. Пусть случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ имеют равномерные на отрезке $[0, 1]$ распределения. Показать, что распределение величины $\eta = \prod\limits_{i=1}^{n} \xi_i$ имеет вид

$$P\{\eta < s\} = \frac{(-1)^n}{n!} \int\limits_0^s t(\ln t)^n \, dt.$$

У к а з а н и е. Либо вычислить непосредственно, либо ввести новые переменные $\xi_i = \exp(-\xi_i)$.

13. Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — n независимых равномерно распределенных на $[0, t]$ случайных величин. Показать, что среднее значение их минимума равно $1+\frac{t}{n}$, а дисперсия $nt^2[(n+1)^2(n+2)]^{-1}$.

14. Число происшествий за неделю на некотором производстве является случайной величиной со средним μ и дисперсией σ^2. Количество травм, полученных в результате различных происшествий, представляет собой независимые величины с одинаковым средним ν и дисперсиями τ^2. Показать, что среднее число травм за неделю равно $\mu\nu$, а дисперсия числа травм за неделю равна $\nu^2\sigma^2 + \mu\tau^2$.

15. Пусть ξ_1 и ξ_2 — независимые случайные величины, равномерно распределенные на отрезке $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$. Показать,что разность $\xi_1 - \xi_2$ имеет распределение, не зависящее от θ, и найти ее плотность вероятности.

16. Случайные величины ξ и η обладают следующими свойствами: ξ — положительна, то есть $P\{\xi > 0\} = 1$, и имеет непрерывную плотность $f(x)$; η при фиксированном ξ имеет равномерное распределение на отрезке $[0, \xi]$. Доказать, что если η и $\xi - \eta$ независимы, то

$$f(x) = a^2 x\ e^{-ax},\ x > 0,\ a > 0.$$

2 Лабораторная работа №2 (2 часа)

ТЕМА: *Характеристические функции.*

Необходимые понятия и теоремы: характеристические функции и их свойства, прямая и обратная предельные теоремы.

Литература: [1] стр.149-155, 178-185; [2] стр.199-272; [3] стр. 52-56; [5] стр. 129-154, 174-188; [7] стр.292-316, 342-350;

Задание 2.1 *Найти характеристическую функцию случайной величины, имеющей:*

1) распределение Бернулли;

2) биномиальное распределение;

3) распределение Пуассона;

4) геометрическое распределение;

5) равномерное на отрезке $[a, b]$ распределение;

6) нормальное (Гаусса) распределение;

7) показательное (экспоненциальное) распределение;

8) распределение Коши.

Задание 2.2 *Найти распределения, которым соответствуют следующие характеристические функции:*

1)	$\cos t$;	*5)*	$1/(1+t^2)$;
2)	$\cos^2 t$;	*6)*	$1/(1-it)$;
3)	$\exp(-t^2)$;	*7)*	$(\sin t)/t$;
4)	$\exp(-\|t\|)$;	*8)*	$\exp(-\|t\|)\cos t$.

Задание 2.3 *Могут ли следующие функции быть характеристическими функциями?*

1)	$\sin t$;	*5)*	$\exp(-i\|t\|)$;
2)	$\sin(t)+1$;	*6)*	$\cos t^2$;
3)	$1/(1+t^4)$;	*7)*	$\exp(-t^4)$;
4)	$\|\cos t\|$;	*8)*	$(\sin t)/t$.

Задачи.

1. Доказать, что характеристическая функция вещественна тогда и только тогда, когда она четна.

2. Доказать, что четная характеристическая функция представима в виде:

$$f(t) = \int_{-\infty}^{+\infty} \cos tx \; dF(x).$$

3. Пусть ξ и η независимые одинаково распределенные случайные величины с характеристической функцией $f(t)$. Найти характеристическую функцию $\xi - \eta$.

4. Случайные величины ξ_1 и ξ_2 независимы имеют вид:

$$\begin{array}{cccc} \xi_i: & 0 & 1 & 2 \\ p: & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \qquad i = 1, 2.$$

Найти характеристическую функцию случайной величины $\xi = \min(\xi_1, \xi_2)$.

5. Случайные величины ξ_1, ξ_2 и $\varkappa$ — независимы. Характеристические функции ξ_1 и ξ_2 равны соответственно $f_1(t)$ и $f_2(t)$, $P(\varkappa = 1) = 1 - P(\varkappa = 0) = p \in (0, 1)$. Найти характеристическую функцию случайной величины $\eta = \varkappa\xi_1 + (1 - \varkappa)\xi_2$.

6. Доказать, что характеристическая функция четна тогда и только тогда, когда соответствующая функция распределения $F(x)$ удовлетворяет соотношению $F(x) = 1 - F(-x - 0)$.

7. Доказать, что $f_\xi(t) \geq 1 - \sigma^2 t^2/2$, если характеристическая функция $f_\xi(t)$ — вещественна, $M\xi = 0$, $D\xi = \sigma^2$.

8. Случайная величина ξ принимает значения 1 и -1 с вероятностью 1/2 каждое. Вычислить характеристическую функцию ξ.

9. Доказать, что функция $\cos^n t$ является характеристической и найти соответствующее распределение вероятностей.

10. Распределение случайной величины ξ называют решетчатым с шагом решетки h, если существуют a и h такие, что

$$\sum_{k=-\infty}^{+\infty} P(\xi = a + kh) = 1.$$

Доказать, что случайная величина ξ с характеристической функцией φ_ξ имеет решетчатое распределение с шагом h тогда и только тогда, когда

$$|\varphi_\xi(2\pi/h)| = 1.$$

11. Случайная величина ξ имеет Γ-распределение, если ее плотность

$$p(x,\alpha,\lambda) = \begin{cases} \frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

$$\Gamma(\lambda) = \int_0^{+\infty} x^{\lambda-1} e^{-x} dx, \ \lambda > 0, \ \alpha > 0.$$

Найти характеристическую функцию ξ.

12. Найти плотность распределения случайной величины

$$\chi_n^2 = \sum_{i=1}^{n} \xi_i^2,$$

используя аппарат характеристических функций, где ξ_i — независимы и нормально распределены с параметрами (0, 1).

3 Лабораторная работа №3 (4 часа)

ТЕМА: *Предельные теоремы теории вероятностей.*

Необходимые понятия и теоремы: центральная предельная теорема для одинаково распределенных слагаемых, теорема Ляпунова, условие Линдеберга, закон больших чисел, усиленный закон больших чисел, теоремы Чебышева, Маркова, Колмогорова.

Литература: [1] стр.149-155, 178-185; [2] стр.199-272; [3] стр. 29-35; [5] стр. 129-154, 174-188; [7] стр.350-357, 376-384;

Задание 3.1 *Случайная величина η является средним арифметическим n независимых одинаково распределенных случайных величин с математическим ожиданием равным m и дисперсией равной d. Найти вероятность того, что η принимает значения в промежутке (k_1, k_2)*

1)	$n = 3200,$	$m = 1,$	$d = 2,$	$k_1 = 0.95$	$k_2 = 1.05;$
2)	$n = 1600,$	$m = 2,$	$d = 4,$	$k_1 = 1.95$	$k_2 = 2.05;$
3)	$n = 1800,$	$m = 1,$	$d = 2,$	$k_1 = 0.95$	$k_2 = 1.05;$
4)	$n = 1200,$	$m = 2,$	$d = 3,$	$k_1 = 1.95$	$k_2 = 2.05;$
5)	$n = 1600,$	$m = 3,$	$d = 4,$	$k_1 = 2.95$	$k_2 = 3.05;$
6)	$n = 800,$	$m = 1,$	$d = 2,$	$k_1 = 0.95$	$k_2 = 1.05;$
7)	$n = 800,$	$m = 3,$	$d = 8,$	$k_1 = 2.95$	$k_2 = 3.05;$
8)	$n = 1600,$	$m = 1,$	$d = 4,$	$k_1 = 0.95$	$k_2 = 1.05.$

Задание 3.2 *Пусть $\xi_1, \xi_2, \ldots \xi_n$, $n \geq 1$ — независимые одинаково распределенные случайные величины. Найти функции распределения случайных величин $\eta_n = \sum\limits_{k=1}^{n} \xi_k$, $\zeta_n = \dfrac{\eta_n}{n}$, $\varkappa_n = \dfrac{\eta_n}{\sqrt{n}}$, $\theta_n = \dfrac{\eta_n - M\eta_n}{\sqrt{n}}$ в том случае, когда ξ_k $(1 \leq k \leq n)$ имеют распределения:*

1) Бернулли;
2) биномиальное;
3) Пуассона;
4) геометрическое;
5) равномерное на $[0, 1]$;
6) нормальное;
7) показательное (экспоненциальное);
8) Коши.

Найти $\lim\limits_{n\to\infty} F_{\theta_n}(x)$, $\lim\limits_{n\to\infty} F_{\varkappa_n}(x)$, $\lim\limits_{n\to\infty} F_{\zeta_n}(x)$, $\lim\limits_{n\to\infty} F_{\eta_n}(x)$, где $F_\xi(x)$ — функция распределения случайной величины ξ.

Задание 3.3 *Установить, будут ли выполнены центральная предельная теорема, закон больших чисел, усиленный закон больших чисел для последовательности независимых случайных величин ξ_n, $n \geq 1$ с распределениями задаваемыми следующим образом:*

1) $P = \{\xi_n = \pm 2^n\} = \frac{1}{2}$;

2) $P = \{\xi_n = \pm 2^n\} = \frac{1}{2^{2n+1}}$, $\quad P = \{\xi_n = 0\} = 1 - \frac{1}{2^{2n}}$;

3) $P = \{\xi_n = \pm n\} = \frac{1}{2\sqrt{n}}$, $\quad P = \{\xi_n = 0\} = 1 - \frac{1}{\sqrt{n}}$;

4) $P = \{\xi_n = \pm\sqrt{n}\} = \frac{1}{n}$ $\quad P = \{\xi_n = 0\} = 1 - \frac{2}{n}$;

5) $P = \{\xi_n = \pm 1\} = \frac{1}{2} - \frac{1}{2^{n+1}}$, $\quad P = \{\xi_n = \pm 2\} = \frac{1}{2^{n+1}}$;

6) $P = \{\xi_n = \pm n\} = \frac{1}{2^n}$, $\quad P = \{\xi_n = 0\} = 1 - \frac{1}{2^{n-1}}$;

7) $P = \{\xi_n = \pm\sqrt{n}\} = \frac{1}{2\sqrt[3/2]{n}}$, $\quad P = \{\xi_n = 0\} = 1 - \frac{1}{\sqrt[3/2]{n}}$;

8) $P = \{\xi_n = \pm\sqrt{n}\} = \frac{1}{2n}$, $\quad P = \{\xi_n = 0\} = 1 - \frac{1}{n}$;

Задачи.

1. Дисперсия каждой из 4500 независимых одинаково распределенных случайных величин равна 5. Найти вероятность того, что среднее арифметическое этих случайных величин отклонится от своего математического ожидания не более чем на 0.04.

2. Случайная величина ξ_λ имеет распределение Пуассона с параметром $\lambda > 0$. Доказать, что при $\lambda \to \infty$ случайная величина

$$\frac{\xi_\lambda - \lambda}{\sqrt{\lambda}}$$

асимптотически нормальна с параметрами $(0,1)$.

3. Случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ независимы, $D\xi_n = n^\alpha$, $0 < \alpha < 1$. Применим ли к данной последовательности закон больших чисел?

4. Пусть ξ_n, $n \geq 1$ — последовательность случайных величин, такая, что ξ_n может зависеть только от ξ_{n-1} и ξ_{n+1}, но не зависит от всех других ξ_k. Выполняется ли для нее закон больших чисел, если $D\xi_n \leq c < \infty$, $n = 1, 2, \ldots$?

5. Дана последовательность независимых случайных величин $\{\xi_n\}_{n=1}^{\infty}$, распределенных по нормальному закону с параметрами $(0, \sigma_n^2)$. Найти закон распределения $\xi_1 + \xi_2 + \ldots + \xi_n$ при $n \to \infty$, если

$$\sum_{k=1}^{\infty} \sigma_k^2 = \sigma^2 < \infty.$$

6. Дана последовательность независимых случайных величин $\xi_1, \ldots, \xi_n, \ldots$, равномерно распределенных на отрезке $[-a_n, a_n]$. Применима ли к данной последовательности центральная предельная теорема, если $a_n \geq c > 0$, $n = 1, 2, \ldots$?

7. Найти

см Севастьянов стр. 152

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}.$$

8. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — последовательность случайных величин, для которой выполняется закон больших чисел. Обязан ли выполняться закон больших чисел для последовательности $|\xi_1|, |\xi_2|, \ldots, |\xi_n|, \ldots$?

9. Независимые случайные величины $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ могут принимать только по два равновероятных значения α_1 и $-\alpha_1$, $2\alpha_2$ и $-2\alpha_2, \ldots$, $n\alpha_n$ и $-n\alpha_n, \ldots$; $\alpha_n > \alpha > 0$, $n = 1, 2 \ldots$. Применим ли к этой последовательности закон больших чисел?

10. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ последовательность независимых одинаково распределенных случайных величин с конечными дисперсиями. Доказать, что для любого вещественного числа x предел

$$\lim_{n\to\infty} P\{\xi_1 + \xi_2 + \ldots \xi_n < x\}$$

равен либо 0, либо 1, либо 1/2.

Указать условия, при которых имеет место каждая из указанных ситуаций.

11. Пусть $P_n = \max_{0\leq k\leq n} P\{\mu_n = k\}$, где μ_n – число успехов в схеме Бернулли, для которой вероятность успеха в каждом независимом испытании равна p. Найти $\lim\limits_{n\to\infty} \sqrt{n}\cdot P_n$.

12. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ последовательность независимых одинаково распределенных случайных величин с ненулевыми математическими ожиданиями и конечными дисперсиями, $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Найти $D\xi_i$, если предел

$$\lim_{n\to\infty} P\left\{\frac{\eta_n}{\sqrt{n}} > 1\right\} = \frac{1}{3}.$$

13. Случайные величны ξ_k, $k = 1, 2, \ldots$ независимы, причем случайные величины ξ_k с четными номерами k имеют распределение Бернулли с параметром p, $0 < p < 1$, а с нечетными номерами — нормальное распределение $\mathcal{N}(a, \sigma^2)$. Найти

$$\lim_{n\to\infty} F_{\zeta_n}(x),\ \zeta_n = \frac{\alpha \sum\limits_{k=1}^{n} (\xi_k - M\xi_k)}{\sqrt{\sum\limits_{k=1}^{n} D\xi_k}},\ \alpha \neq 0.$$

14. Доказать, что условие Линдеберга выполнено для последовательности независимых одинаково распределенных случайных величин с конечной дисперсией.

15. Доказать, что если для последовательности случайных величин выполнено условие Ляпунова, то выполнено и условие Линдеберга.

16. Последовательность независимых случайных величин $\xi_1, \ldots, \xi_n, \ldots$ называется эквивалентной последовательности случайных величин $\eta_1, \eta_2, \ldots, \eta_n, \ldots$, если сходится ряд

$$\sum_{n=1}^{\infty} P(\omega : \xi_n(\omega) \neq \eta_n(\omega)).$$

Доказать, что если при всех $n \in \mathbb{N}$ случайные величины ξ_n и η_n имеют одинаковые математические ожидания и закон больших чисел применим к одной из двух эквивалентных последовательностей, то он применим и к другой.

17. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ последовательность независимых одинаково распределенных невырожденных случайных величин с конечными дисперсиями, $\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n$. Доказать, что для любых конечных вещественных чисел a и b

$$\lim_{n\to\infty} P\{a \le \eta_n \le b\} = 0.$$

18. Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ последовательность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием, единичной дисперсией и абсолютно интегрируемой характеристической функцией. Пусть $p_n(x)$ — плотность распределения случайной величины

$$\eta_n = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{\sqrt{n}}.$$

Доказать, что равномерно по x

$$\lim_{n\to\infty} p_n(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}.$$

19. Пусть функция $g(x), x \geq 0$ неотрицательна и монотонно возрастает. Доказать, что для любой действительной случайной величины ξ справедливо неравенство

$$P\{|\xi| \geq x\} \leq \frac{Mg(|\xi|)}{g(x)}.$$

20. Оценить снизу вероятность того, что при 1000 бросаниях монеты число μ выпадений герба будет заключено между 450 и 550.

21. Пусть заданы m последовательностей случайных величин $\xi_n^{(k)}, k = 1, 2, \ldots, m$, $\xi_n^{(k)} \xrightarrow{P} \xi^{(k)}$, $n \to \infty$, а $\Phi(x_1, x_2, \ldots, x_m)$ — непрерывная функция аргументов $x_1, x_2, \ldots, x_m$, определенная на $\mathbb{R}^m$. Тогда

$$\Phi(\xi_n^{(1)}, \xi_n^{(2)}, \ldots, \xi_n^{(m)}) \xrightarrow{P} \Phi(\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(m)}), \; n \to \infty.$$

22. Пусть $\{\xi_n\}_{n=1}^{\infty}$ — последовательность независимых равномерно распределенных на отрезке $[0, 1]$ случайных величин. Доказать, что

$$\eta_n = \left(e^n \prod_{k=1}^{n} \xi_k\right)^{1/n} \xrightarrow{P} 1, \; n \to \infty.$$

23. Пусть $\{\xi_n\}_{n=1}^{\infty}$ — последовательность независимых одинаково распределенных случайных величин, $M\xi_n = a$, $D\xi_n = \sigma^2$, $P\{\omega \,|\, \xi_1(\omega) = 0\} = 0$. Доказать, что последовательность случайных величин

$$\eta_n = \frac{\xi_1 + \ldots + \xi_n}{\xi_1^2 + \ldots + \xi_n^2}$$

сходится по вероятности и вычислить предел.

24. Пусть случайная величина η_n равна сумме очков, появившихся при n бросаниях симметричной игральной кости. Используя неравенство Чебышева, оценить сверху

$$P\left\{\left|\frac{\eta_n}{n} - 3.5\right| > \varepsilon\right\}.$$

4 Лабораторная работа №4 (2 часа)

ТЕМА: *Цепи Маркова. Основные понятия теории случайных процессов.*

Необходимые понятия и теоремы: простая цепь Маркова, переходная вероятность, начальная вероятность, однородная цепь Маркова, уравнение Маркова, характеристическая матрица, несущественное, существенное состояния, связанное состояние, возвратное, невозвратное состояния, периодическое, эргодическое состояния; случайный процесс, одномерный закон распределения случайного процесса, математическое ожидание и дисперсия случайного процесса, корреляционная функция, непрерывный случайный процесс, стационарный случайный процесс.

Литература: [1] стр.270-292, 321-345; [2] стр.109-125, 238-338; [3] стр.109-136, 179-194; [5] стр. 77-84; [7] стр.121-144, 390-457.

Задачи.

1. Вероятности перехода в простой однородной цепи Маркова даются матрицей

$$\pi = \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \\ 1/2 & 1/3 & 1/6 \end{pmatrix}$$

а) Чему равно число состояний этой цепи.

б) Найти вероятности перехода из состояния в состояние за два шага.

2. Электрон может находиться на одной из счетного множества орбит в зависимости от наличной энергии. Переход с i-ой на j-ю орбиту происходит за одну секунду с вероятностью

$$c_i e^{-\alpha|i-j|} \ (\alpha > 0).$$

Найти: а) вероятности прехода за 2 секунды; б) постоянные c_i.

3. Имеется простая однородная цепь Маркова с матрицей перехода $\pi = (p_{ij})$. Вычислить:

а) вероятность состояния E_i на n-ом шаге, если известны все последующие состояния системы;

б) вероятность E_i на n-ом шаге, если известны все предыдущие и последующие состояния.

4. Точка движется по прямой и в течение каждой очередной секунды движения может или сместиться на единицу расстояния, или остаться на месте.

Заданы: 1) вероятность p_1 смещения для первой секунды; 2) вероятность α ($\alpha = const$) смещения для любой рассматриваемой секунды, если известно, что в предыдущюу секунду также произошло смещение; 3) вероятность β ($\beta = const$) смещения для любой рассматриваемой секунды, если известно, что в течение предыдущей секунды точка оставалась на месте.

Найти вероятность смещения токи в течение $n+1$-й секунды.

5. Цепь Маркова управляется матрицей перехода

$$\pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Эргодична ли эта цепь?

6. Вероятность перехода дается матрицей

$$\pi = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

а) Убедиться в эргодичности этой цепи.

б) Найти предельные вероятности.

7. Простая однородная цепь Маркова с двумя состояниями имеет матрицу перехода

$$\pi = \begin{pmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{pmatrix}, \; p_1 \neq 0, p_2 \neq 0.$$

а) Составить характеристическое уравнение и найти характеристические числа матрицы.

б) Найти финальные вероятности $\overline{p_1}$ и $\overline{p_2}$.

8. Простая однородная цепь Маркова с двумя состояниями управляется матрицей

$$\pi = \begin{pmatrix} 1 & 0 \\ 1/3 & 2/3 \end{pmatrix}.$$

Убедиться в регулярности этой цепи и найти финальные вероятности.

9. В некоторой области пространства имеются однородные частицы. Под влиянием случайных причин в эту область извне могут проникать новые такие же частицы (или там могут возникать новые частицы), но они не могут покидать этой области (или исчезать). Если в момент t в области имеется n частиц, то вероятность (условная) того, что за промежуток $(t, t+\Delta t)$ в этой области появится новая частица, не зависит от t и равна $\lambda_n \Delta t + o(\Delta t)$, где λ_n — неотрицательная постоянная. Вероятность того, что за промежуток времени длиной Δt в области появятся две или более новых частиц, равна $o(\Delta t)$. В начальный момент $t = 0$ в области имелось 0 частиц.

а) Составить систему дифференциальных уравнений, определяющих (вместе с начальными условиями) вероятности $p_n(t)$ того, что в момент t (любой) в области будет ровно n частиц.

б) Найти явные выражения вероятностей $p_n(t)$.

в) Положив $\lambda_n = 3^n$, $n = 0, 1, 2, \ldots$, исследовать при больших значениях t сумму $\sum\limits_{n=0}^{\infty} p_n(t)$. Каков вероятностный смысл того, что при достаточно больших t эта сумма меньше единицы?

10. В условиях предыдущей задачи доказать, что если $\lambda_n > 0$ при всех n, то $p_n(t)$ при $n \geq 1$ имеет единственный максимум в некоторой точке и стремится к нулю при $t \to \infty$.

Показать, что $t_1 < t_2 < t_3 < \ldots$ (использовать метод индукции и продифференцировать дифференциальные уравнения процесса).

11. Доказать, что в условиях предыдущей задачи при $n \to \infty$ $t_n \to \infty$, если процесс чистого размножения не является расходящимся.

12. Процесс Пуассона. Пусть некоторая физическая система может находиться в одном из счетного множества состояний $E_1, E_2, E_3, \ldots$, причем в любой момент времени t она может сменить свое состояние, перейдя в состояние с номером, на единицу большим. Если Δt достаточно мало, то вероятность перехода системы за промежуток времени $(t, t + \Delta t)$ из состояния E_n в E_{n+1} равна $\lambda\Delta t + o(\Delta t)$, где λ — положительная постоянная. Составить системы дифференциальных уравнений, описывающих этот процесс, и, решив ее, найти вероятности $p_n(t)$ $(n = 1, 2, \ldots)$ того, что система в момент времени t окажется в состоянии E_n, если известно, что в начальный момент $t_0 = 0$ система находилась в состоянии E_0.

13. Процесс Юла. В области G имеются частицы, способные размножаться (путем деления или иначе) и остающиеся в этой области в дальнейшем. За малый промежуток времени $(t, t + \Delta t)$ каждая частица с вероятностью $\lambda\Delta t + o(\Delta t)$ производит новую частицу независимо от остальных частиц.

а) Составить систему дифференциальных уравнений, определяющих этот процесс.

б) Решить эту систему.

в) Найти математическое ожидание и дисперсию распределения, определяемого этой системой.

14. Процесс чистой гибели. В облсти G в начальный момент времени $t = 0$ находилось k частиц. Независимо друг от друга частицы могут исчезать из этой области, причем каждая частица за малый промежуток времени Δt может исчезнуть с вероятностью $\lambda\Delta t + o(\Delta t)$. Новые частицы в области появиться не могут.

а) Найти дифференциальные уравнения, описывающие этот процесс.

б) Найти вероятности $p_n(t)$ — решения этих уравнений.

в) Найти математическое ожидание и дисперсию числа часиц в области G к моменту t.

15. Процесс Пойа. В области G появляются некоторые частицы и в дальнейшем остаются в этой области. Если к моменту t в области имелось n частиц, то вероятность (условная) увеличения их числа на единицу за малый промежуток времени $(t, t + \Delta t)$ равна

$$\frac{1 + an}{1 + at}\Delta t + o(\Delta t),$$

где a — некоторая положительная постоянная. Вероятность увеличения числа частиц за то же время на две или более равна $o(\Delta t)$.

а) Составить дифференциальные уравнения, определяющие этот процесс (то есть вероятности $p_n(t)$ того, что к моменту t в области G будет n частиц).

б) Найти явные выражения $p_n(t)$.

в) Найти математическое ожидание и дисперсию числа частиц в области.

16. Процесс размножения и гибели. В области G имеются однородные частицы (например, бактерии), которые могут порождать такие же новые частицы, например, посредством деления на две, а также могут погибать (исчезать). Если

Δt мало, то вероятность для каждой частицы (независимо от наличия и поведения остальных частиц) породить одну новую равна $\lambda \Delta t + o(\Delta t)$, а вероятность погибнуть равна $\mu \Delta t + o(\Delta t)$, где λ и μ — некоторые положительные постоянные.

а) Составить систему обыкновенных дифференциальных уравнений, определяющую этот процесс.

б) При условии, что $p_n(t)$ $(n = 0, 1, 2 \ldots)$ — решение этой системы, то есть вероятность того, что к моменту t в области G окажется n частиц, составить дифференциальное уравнение в частных производных, которому удовлетворяет производящая функция этих вероятностей, то есть функция

$$F(x,t) = \sum_{n=0}^{\infty} p_n(t) x^n.$$

в) Считая, что в начальный момент $t = 0$ в области G находилась одна частица, решить это дифференциальное уравниние и найти явное выражение производящей функции $F(x,t)$.

5 Лабораторная работа №5 (4 часа)

ТЕМА: *Основные понятия и элементы выборочной теории. Оценивание неизвестных параметров распределения.*

Необходимые понятия и теоремы: вариационный ряд выборки, гистограмма, выборочные характеристики, распределения χ^2, Стьюдента, Фишера, несмещенность, состоятельность, достаточные статистики, методы моментов и максимального правдоподобия, доверительные интервалы и доверительные вероятности.

Литература: [2] стр.127-196; [3] стр.234-266; [5] стр.189-194, 213-236; [13] стр.13-95; [14] стр.15-31, 51-69, 198-211.

Задание 5.1 *Найти плотности*

1) распределения хи-квадрат с n степенями свободы;

ШИРЯЕВ, 159

2) распределения Фишера со степенями свободы m и n;

3) распределения Стьюдента с m степенями свободы.

Задание 5.2 *Дана реализация выборки*

1) 0.30 -0.01 -0.13 -0.17 -0.10 0.41 0.40 0.44 -0.10
-0.02 -0.21 0.01 0.20 0.26 -0.08 -0.15 0.01 0.17
0.18 -0.15 -0.14 0.06 0.25 0.02 -0.05 0.38 -0.11
0.01 -0.08 -0.03 0.01 -0.07 0.02 0.10 0.10 -0.07
0.05 0.08 -0.12 -0.28 0.08 0.08 -0.05 -0.09 -0.15
-0.03 -0.06

2) 0.32 0.02 -0.18 0.16 -0.13 -0.02 0.08 -0.04 -0.04
0.06 0.21 0.39 -0.12 -0.01 -0.02 0.18 0.16 0.03
0.09 0.09 0.26 -0.01 -0.03 0.13 -0.12 0.01 -0.03
0.07 -0.05 0.04 0.03 0.10 0.09 -0.05 -0.01 0.20
-0.06 0.03 0.05 0.22 -0.20 0.27 -0.06 -0.03 0.03
-0.04 -0.04 0.06

3) 0.32 0.08 -0.23 -0.01 0.09 0.13 -0.05 -0.05 0.05
-0.03 0.02 0.06 -0.02 0.09 -0.12 0.04 -0.01 0.22
-0.04 -0.18 0.21 0.18 0.26 0.01 0.03 0.20 -0.20
0.03 -0.13 -0.04 -0.12 0.03 -0.03 0.07 0.09 0.03
-0.06 0.06 -0.02 0.16 -0.01 -0.03 0.10 -0.06 0.25
-0.04 -0.11 0.05

4) 0.28 -0.02 -0.11 -0.12 0.38 0.36 -0.16 0.38 -0.08
-0.03 -0.20 0.01 0.18 0.28 -0.06 -0.14 0.01 0.16
0.18 -0.14 0.07 0.24 0.03 -0.07 -0.13 0.35 -0.12
0.01 -0.09 -0.04 0.02 -0.08 0.03 0.09 0.10 -0.06
0.04 0.07 -0.12 -0.24 0.09 0.08 -0.06 -0.08 -0.14
-0.04 -0.06

5) 0.30 0.03 -0.16 0.18 -0.12 -0.02 0.08 -0.05 -0.05
0.05 0.22 0.32 -0.13 -0.01 -0.03 0.18 0.15 0.03
0.10 0.08 0.28 -0.02 -0.03 0.14 -0.12 0.01 -0.04
0.05 -0.07 0.03 0.08 0.09 -0.04 -0.03 -0.02 0.21
-0.06 0.03 0.04 -0.22 0.20 0.26 -0.05 -0.03 0.03
-0.06 -0.02 0.06

6) 0.32 0.10 -0.24 -0.01 0.08 0.14 -0.06 -0.04 0.05
-0.01 0.01 0.07 -0.02 0.09 -0.11 0.05 -0.01 0.23
-0.04 -0.17 0.23 0.16 0.25 0.01 0.03 0.19 -0.19
0.03 -0.13 -0.03 -0.12 0.03 -0.02 0.05 0.11 0.04
-0.07 0.07 -0.01 0.16 -0.02 -0.03 0.13 -0.08 0.24
-0.04 -0.11 0.08

7) 0.32 -0.03 -0.12 -0.14 0.34 0.36 -0.15 0.32 -0.10
-0.04 -0.18 0.02 0.17 0.26 -0.05 -0.13 0.01 0.15
0.17 -0.13 0.08 0.23 0.02 -0.06 -0.12 0.31 -0.14
0.02 -0.10 -0.03 0.01 -0.08 0.04 0.07 0.10 -0.05
0.03 0.08 -0.14 -0.23 0.08 0.07 -0.06 -0.06 -0.16
-0.05 -0.07

8) 0.32 -0.03 -0.12 -0.14 0.34 0.36 -0.15 0.32 -0.10
0.25 0.04 -0.13 -0.17 -0.11 -0.01 0.06 -0.04 -0.06
0.04 0.20 0.28 -0.12 0.01 -0.05 0.19 0.14 0.02
0.11 0.08 0.20 -0.02 -0.03 0.13 -0.11 0.02 -0.04
0.04 -0.08 0.04 0.06 0.11 -0.03 0.01 -0.05 0.10
0.06 -0.02 -0.05

1. Сгруппировать выборочные данные.

У к а з а н и е. *Группировку выборочных данных проводят в следующей последовательности:*

а) строят вариационный ряд данной реализации

$$x_1 \geq x_2 \geq \ldots \geq x_n;$$

б) по вариационному ряду, используя правило Старджеса

$$\Delta t = \frac{x_1 - x_n}{1 + 3.3 \lg n},$$

находят оптимальную величину интервала Δt;

в) составляют таблицу 1 по следующей форме:

№	Границы интервала		Середина	Частота	Частость
п/п	Начало $t_i^{н}$	Конец $t_i^{к}$	интервала t_i	n_i	$n_i^{o} = \dfrac{n_i}{n}$
1					
⋮					
N					

г) объединяя соседние интервалы так, чтобы для любого $i = \overline{1, N}$, $n_i \geq 5$ *приходят к таблице 2.*

№	Границы интервала		Середина	Частота	Частость
п/п	Начало $t_i^{н*}$	Конец $t_i^{к*}$	интервала t_i^*	n_i^*	n_i^{o*}
1					
⋮					
N^*					

2. Найти a *и* σ^2 *по формулам:*

$$a = \frac{1}{n} \sum_{i=1}^{N^*} t_i^* n_i^*,$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{N^*} (t_i^* - a)^2 n_i^*.$$

Выяснить, как согласуются a *и* σ^2*, соответственно, с выборочным средним и выборочной дисперсией.*

3. Построить на одном чертеже гистограмму и плотность распределения

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-a)^2}{2\sigma^2}}.$$

4. Найти доверительные интервалы для среднего и дисперсии в модели $\mathcal{N}(a, \sigma^2)$, если доверительная вероятность равна 0.95.

Задание 5.3 *Найти достаточную статистику для*

1) θ *в модели* $R(0, \theta)$;
2) $\theta = (\theta_1, \theta_2)$ *в модели* $R(\theta_1, \theta_2)$;
3) θ *в модели* $K(\theta)$;
4) θ *в модели* $Bi(1, \theta)$;
5) θ *в модели* $\Pi(\theta)$;
6) θ *в модели* $\mathcal{N}(\theta, \sigma^2)$;
7) θ *в модели* $\mathcal{N}(\mu, \theta^2)$;
8) $\theta = (\theta_1, \theta_2)$ *в модели* $\mathcal{N}(\theta_1, \theta_2^2)$.

Задание 5.4 *Методом моментов и методом максимального правдоподобия оценить неизвестный параметр θ (векторный или скалярный) в модели*

1) $R(0, \theta)$; *2)* $R(\theta_1, \theta_2)$; *3)* $\Gamma(\theta, 1)$; *4)* $Bi(1, \theta)$;

5) $\Pi(\theta)$; *6)* $\mathcal{N}(\theta, \sigma^2)$; *7)* $\mathcal{N}(\mu, \theta^2)$; *8)* $\mathcal{N}(\theta_1, \theta_2^2)$.

Выяснить, являются ли полученные оценки несмещенными, состоятельными.

Задачи.

1. Цена на некоторый товар в 15 торговых точках составила: 422, 419, 426, 420, 426, 423, 432, 428, 438, 434, 411, 417, 414, 441,

420. Построить график эмпирической функции распределения и гистограмму; в предположении гауссовской модели данных построить несмещенные оценки математического ожидания (средней цены) μ и дисперсии цены σ^2.

2. Уровень воды в реке по отношению к номиналу измеряется в течение 44 весенних паводков, и данные измерений приведены в следующей таблице:

Уровень в см	0-24	25-49	50-74	75-99	100-124	125-149
Число случаев	0	1	3	6	7	6
Уровень в см	150-174	175-199	200-299	300-399	> 400	
Число случаев	5	4	8	4	0	

a) Построить гистограмму частот и эмпирическую функцию распределения уровня воды ξ в реке.

b) Вычислить выборочное среднее и дисперсию, а также выборочные коэффициенты асимметрии и эксцесса.

c) Принимая во внимание

$$p_\xi(x) = \frac{k^{\alpha-1}\ x^\alpha\ e^{-kx}}{\Gamma(\alpha+1)},\ k = 0.024, \alpha = 3,$$

найти ε из условия $P\{|\overline{x} - \mu| \le \varepsilon\} = 0.95$ и сравнить с ε отклонение выборочного среднего от теоретического μ.

3. Вывести следующие формулы для выборочных моментов A_{nk} k – ого порядка и выборочной дисперсии S^2:

$$M\{A_{nk}\} = M\xi^k = \alpha_k,\ cov\{A_{nk}, A_{ns}\} = \frac{\alpha_{k+s} - \alpha_k\alpha_s}{n};$$

$$M\{S^2\} = \frac{n-1}{n}\mu_2,\ D\{S^2\} = \frac{(n-1)^2}{n^3}\left(\mu_4 - \frac{n-1}{n-3}\mu_2^2\right),$$

где

$$\mu_k = M(\xi - \alpha_1)^k,\ cov\{\overline{x}, S^2\} = \frac{n-1}{n^2}\mu_3.$$

4. В основе алгоритма моделирования значений случайной величины с пуассоновским распределением $\Pi(\lambda)$ лежит следующий факт: если U_i, $i = 0, 1, 2, \ldots$ — независимые случайные величины с равномерным на отрезке $[0, 1]$ распределением и

$$\xi = \max\left\{k : \prod_{i=0}^{k} U_i \geq e^{-\lambda}\right\},$$

тогда ξ имеет пуассоновское распределение $\Pi(\lambda)$. Доказать этот факт.

5. В ходе маркетинговых исследований автомобильного рынка для анализа экономичности двигателей была зарегистрирована случайная выборка 12 автомобилей одного класса:

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	18.2	19.5	20.7	17.8	19.3	20.5	17.8	19.9	20.3	18.6	19.0	20.8

Здесь x_i — показатель экономичности i-го автомобиля (расстояние, преодолеваемое этим автомобилем при использовании 1 галлона бензина). Используя гауссовскую модель выборки $\mathcal{N}(\theta_1, \theta_2)$, методом максимального правдоподобия построить несмещенные оценки параметров θ_1 (средняя экономичность автомобиля), θ_2 (дисперсия экономичности автомобиля) и исследовать оценки на состоятельность и эффективность.

6. Отдел приемочного контроля комплектующих изделий использует следующую методику оценки доли θ дефектных изделий: в большой партии изделий наудачу вскрывается 10 ящиков, изделия наудачу извлекаются по одному и тестируются до обнаружения первого дефектного изделия. Результаты произведенного контроля представлены в таблице:

i	1	2	3	4	5	6	7	8	9	10
x_i	21	25	17	22	19	23	27	16	20	24

Здесь x_i — число годных изделий, извлеченных из i-го ящика до обнаружения первого дефектного изделия. Используя модель выборки из геометрического распределения, найти минимальную достаточную статистику, найти оценку максимального правдоподобия и исследовать ее свойства.

7. При выборочной проверке каждого из 10 малых предприятий налоговой инспекцией не было обнаружено ни одного нарушения Налогового законодательства. Построить 80% –ный, 90% –ный и 99% –ный доверительные интервалы для доли нарушителей Налогового законодательства, используя биномиальную модель распределения для числа нарушителей.

8. Из 30 созданных малых предприятий 8 прекратили свое существование в течение года. Определить интервальную оценку вероятности выживаемости малого предприятия (доверительный уровень 0.95) в предположении, что процесс гибели малых предприятий описывается схемой независимых испытаний Бернулли.

9. При проведении маркетинговых исследований возникла необходимость определения средней цены на некоторый товар по результатам $x_1, x_2, \ldots, x_n$ наблюдений за ценой в течение n месяцев. В предположении гауссовской модели (плотность распределения случайной величины x_i есть нормальное распределение с параметрами $(\theta, 25)$) определить, сколько нужно выполнить наблюдений n, чтобы оценить θ с ошибкой не более 5, при доверительной вероятности 0.95.

10. Владелец ресторана хотел бы знать с вероятностью 0.95, в каких пределах заключена сумма денег, которую посетитель платит за заказ. Случайная выборка объема $n = 36$ обладает выборочным средним 6.5 и выборочным стандартным отклонением 3.

а) Какой должна быть по объему выборка, чтобы абсолютное отклонение средней суммы денег от выборочной средней не превышало 10% с вероятностью 0.95?

б) Найти доверительный интервал с доверительным уровнем 0.99.

11. Число посетителей коммерческого банка в течение рабочего дня является случайной величиной с распределением Пуассона $\Pi(\theta)$, где $\theta > 0$ — неизвестное среднее число посетителей. Данные по числу посетителей в каждый из n дней представлены выборкой $X = \{x_1, x_2, \ldots, x_n\}$. По заданному доверительному уровню $\varepsilon \in (0,1)$ построить "точный" и асимптотически наикратчайший доверительные интервалы.

12. Пусть некоторая фирма, занимающаяся продажей цветов, реализовала в первые $n = 5$ дней своей деятельности $x_1, x_2, \ldots, x_n$ цветов:

i	1	2	3	4	5
x_i	125	89	136	72	104

Чтобы продавать только свежие цветы, директор фирмы желает по имеющимся данным спроса за n дней оценить спрос x_{n+1} на следующий $(n+1)$ -й день с помощью 95% -го доверительного интервала. Показать, используя гауссовскую модель, что с вероятностью 0.95 x лежит в интервале

$$\left(\overline{x} - t_{n-1}(0.975)\ S\sqrt{(n+1)/n}, \overline{x} + t_{n-1}(0.975)\ S\sqrt{(n+1)/n}\right),$$

где $\overline{x}$ — выборочное среднее, S^2 — выборочная дисперсия (несмещенная оценка), $t_n(p)$ — квантиль уровня p для t – распределения Стьюдента с n степенями свободы.

6 Лабораторная работа №6 (4 часа)

ТЕМА: *Проверка статистических гипотез.*

Необходимые понятия и теоремы: статистическая гипотеза и статистическая проверка гипотез, критерий согласия χ^2, критерий однородности χ^2, критерий независимости χ^2.

Литература: [3] стр.299-301; [5] стр.206-211; [13] стр.102-136; [14] стр.211-216.

Задание 6.1 *Методом χ^2 К.Пирсона сравнить эмпирические распределения с теоретическими (см. задание 2 лабораторной работы № 5).*

Задание 6.2 *Поступающие в университет абитуриенты разбиты на два потока по 300 человек в каждом. итоги экзамена по одному и тому же предмету на каждом потоке оказались следующими:*

№ п/п	1 поток "2"	"3"	"4"	"5"	2 поток "2"	"3"	"4"	"5"
1.	33	43	80	144	39	35	72	154
2.	32	44	79	145	40	34	73	153
3.	35	42	79	144	37	36	73	154
4.	34	44	81	141	41	36	73	151

№ п/п	1 поток "2"	"3"	"4"	"5"	2 поток "2"	"3"	"4"	"5"
5.	31	42	79	148	39	35	72	154
6.	33	42	79	146	39	34	71	156
7.	32	42	82	144	40	36	70	154
8.	31	45	78	146	37	37	74	152

Можно ли считать оба потока однородными (на уровне значимости 0.05)?

Задание 6.3 *Для заданной таблицы сопряженности двух признаков проверить гипотезу независимости (уровень значимости взять равным 0.05).*

1.	18	30	2.	17	31	3.	16	32	4.	19	29
	79	173		78	174		77	175		80	172

5.	20	28	6.	17	31	7.	16	32	8.	19	29
	78	174		80	172		81	171		78	174

Задачи.

1. Бытует мнение, что при рождении ребенка вероятность мальчика такая же, как и девочки. Примем это за гипотезу. Для ее проверки имеется огромный статистический материал. Воспользуемся данными по Швейцарии с 1871 по 1900 г.г., когда там родилось $n = 2644757$ человек и среди них $n_1 = 1359671$ мальчиков и $n_2 = 1285086$ девочек. Согласуется ли гипотеза о равновероятности рождения мальчика и девочки с этими числами?

2. Среди 300 человек, поступивших в ВУЗ, 97 имели оценку "5" в школе и 48 получили "5" на вступительных экзаменах по тому же предмету, причем лишь 18 имели "5" и в школе, и на экзаменах. С уровнем значимости 0.1 проверить гипотезу о независимости оценок "5" в школе и на вступительных экзаменах.

3. В следующей таблице приведены данные об урожайности пшеницы на 8 опытных участках одинакового размера:

Номер участка	1	2	3	4	5	6	7	8
Урожайность, ц/га	26.5	25.2	28.9	30.1	32.3	29.3	26.1	25.0

Предполагая, что урожайность — нормально распределенная случайная величина, требуется:

а) проверить гипотезу $H_0 : \mu = 28$ ц/га против альтернативы $H_1 : \mu = 30$ ц/га, $\varepsilon = 0.1$;

б) определить мощность критерия проверки гипотезы H_0;

в) построить график зависимоти мощности — кривую $W = W(\mu)$ (кривую эффективности) критерия.

4. Во время экзамена студентам были предложены задачи из 7 разделов изученного курса. Результаты экзаменов приведены в таблице:

Номер раздела курса	1	2	3	4	5	6	7
Число предложенных задач	165	66	270	160	80	350	150
Доля решенных задач	0.86	0.51	0.52	0.48	0.86	0.41	0.42

Требуется на уровне значимости 0.1 проверить гипотезу о том, что вероятность решения студентами задачи не зависит от раздела, из которого она предложена.

5. Четверо студентов проходят серию из трех тестов, результаты которых следующие:

Студент / Тест	1	2	3
1	12	12	13
2	9	12	12
3	10	11	11
4	10	11	12

Можно ли на уровне значимости 0.05 считать, что студенты имеют одинаковые знания по трем тестовым предметам?

6. Пусть фирма осуществляет массовое производство некоторого изделия и известно, что 96% продукции оказывается годной. Появление дефектных изделий в процессе производства — случайное событие, поэтому отдел контроля качества (ОТК) фирмы проводит выборочную проверку. Известно, что замена негодного изделия годным и его пересылка обходится в \$60, отправка годного изделия — \$20. Как выбрать ОТК критическое значение x показателя качества, чтобы полные расходы фирмы были бы минимальными?

7. Муниципалитет интересуется долей граждан, имеющих телевизор. Пусть по выборке из 500 человек оказалось, что 46% владели телевизорами. Можно ли с вероятностью 0.95 считать, что гипотеза о 50% владении телевизорами является приемлемой?

8. В течение 10 часов регистрировали прибытие автомашин на заправку к бензоколонке. Данные приведены в таблице, где в первой строке указан интервал времени (в часах), во второй — число машин, прибывших в данном интервале:

$a_{i-1} - a_i$	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18
m_i	12	40	22	16	28	6	11	33	18	14

Проверьте на уровне значимости 0.01 гипотезу о том, что время прибытия машин распределено равномерно.

9. При скрещивании двух типов кукурузы во втором поколении обнаружено 4 различных типа растений. Простая менделевская модель предсказывает появление этих четырех типов с вероятностями: 9/16, 3/16, 3/16 и 1/16. В результате наблюдений над 1301 растением получены частоты: $\nu_1 = 773$, $\nu_2 = 231$, $\nu_3 = 238$, $\nu_4 = 59$. При каком уровне значимости ε критерий χ^2 Пирсона подтверждает менделевскую модель?

10. Для проверки подлинности рукописей, писем К.К. Снодграсса, приписывемых М.Твену, Ч.С.Гринегар произвел подсчет числа слов длиной в $k = 1, 2, 3, \ldots 12$ букв в произведениях М.Твена и в письмах К.К.Снодграсса. Выяснилось, что слова из 2, 3 и 4 букв встречаются в произведениях М.Твена с частотами (вероятностями):

k	2	3	4	остальные
Вероятности p_k	0.177	0.232	0.191	0.400

В 13175 словах рукописей писем К.К.Снодграсса слова тех же длин встречаются со следующими частотами:

k	2	3	4	остальные
Абсолютные частоты	2685	2752	2302	5436

С помощью критерия Пирсона на уровне значимости 0.01 проверить гипотезу о том, что письма К.К.Снодграсса представляют собой случайную выборку из произведений М.Твена.

11. Для анализа влияния витамина на профилактику простудных заболеваний 200 человек случайным образом были разделены на две равные группы. Одной группе давали витамин, другой (контрольной) — "пустышку", но всем пациентам было сказано, что им дают витамин. Результаты приведены в таблице:

результат / группа	меньше простудных заболеваний	больше простудных заболеваний	без изменений
контрольная группа	39	21	40
группа, принимавшая витамины	51	20	29
сумма	90	41	69

С помощью критерия χ^2 Пирсона на уровне значимости 0.05 проверить гипотезу о независимости простудных заболеваний от приема витамина.

12. В следующей таблице приведены данные об урожайности кукурузы и риса в различных странах (в ц/га). На уровне значимости 0.05 проверьте гипотезу о том, что средняя урожайность риса и кукурузы одинакова.

Страна	Урожайность	
	Рис	Кукуруза
Колумбия	11.2	19.5
Израиль	40.4	9.7
Аргентина	17.7	33.6
Чили	20.7	26.9
Мексика	9.4	22.5
Тайвань	17.5	32.1
Египет	24.0	52.3
Япония	25.9	50.5

7 Лабораторная работа №7 (2 часа)

ТЕМА: *Линейная регрессия и метод наименьших квадратов.*

Необходимые понятия и теоремы: линейная регресионная модель, метод наименьших квадратов, условное математическое ожидание, уравнение регрессии первого рода, выборочный коэффициен корреляции.

Литература: [3] стр.271-277; [13] стр.179-220.

Задание 7.1 *Пусть зависимость признака y от признака x характеризуется таблицей*

1.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.6	-0.3	-2	-2.3	-1.5	0.7	3.2
2.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.5	-0.8	-1.9	-2.5	-1.4	0.1	3.2

3.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.2	-0.4	-1.8	-2	-1.2	0.2	2
4.	x_i	-3	-2	-1	0	1	2	3
	y_i	1.8	-1.8	-2.2	-2.4	-1.6	1.2	2.4

5.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.4	-1.6	-2.0	-2.8	-1.4	0.8	2.2
6.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.0	-1.2	-2.0	-2.6	-1.6	1.4	2.8

7.	x_i	-3	-2	-1	0	1	2	3
	y_i	2.8	-0.4	-1.8	-2.2	-1.6	0.8	3.0
8.	x_i	-3	-2	-1	0	1	2	3
	y_i	3.0	-0.2	-1.6	-2.0	-1.4	0.6	2.8

Предполагая, что справедлива зависимость

$$y = a_0 + a_1 x + a_2 x^2,$$

определить оценки коэффициентов a_k и оценки среднеквадратичных отклонений: σ — отдельного измерения и σ_{a_k} — коэффициентов a_k. Установить доверительные интервалы для a_k и σ при доверительной вероятности 0.9.

Задание 7.2 *По данной корреляционной таблице найти выборочный коэффициент корреляции и написать выборочные уравнения прямых регрессии η на ξ и ξ на η.*

	y \ x	45	50	55	60	65	70	n_y
	35	3	5	—	—	—	—	8
	35	3	5	—	—	—	—	8
1.	40	—	4	4	—	—	—	8
	45	—	—	7	35	8	—	50
	50	—	—	2	10	8	—	20
	55	—	—	—	5	6	3	14
	n_x	3	9	13	50	22	3	$n = 100$

	y \ x	30	40	50	60	70	80	n_y
	10	2	7	—	—	—	—	9
	20	4	—	5	—	2	—	11
2.	30	—	6	12	—	5	—	23
	40	—	—	8	16	—	8	32
	50	—	—	—	4	10	11	25
	n_x	6	13	25	20	17	19	$n = 100$

	$y \backslash x$	3	8	13	18	23	28	n_y
	4	9	4	1	—	—	—	14
	8	1	10	9	3	—	—	23
3.	12	—	2	6	14	6	—	28
	16	—	—	1	10	—	—	11
	20	—	—	—	—	18	6	24
	n_x	10	16	17	27	24	6	$n = 100$

	$y \backslash x$	20	22	24	26	28	30	n_y
	6	2	2	—	—	—	—	4
	7	2	4	5	6	4	—	21
4.	8	—	2	7	12	10	4	35
	9	—	—	—	10	10	6	26
	10	—	—	—	8	—	6	14
	n_x	4	8	12	36	24	16	$n = 100$

5.	$y \backslash x$	30	35	40	45	50	55	n_y
	12	2	4	—	—	—	—	6
	17	—	6	2	—	—	—	8
	22	—	—	3	50	2	—	55
	27	—	—	1	10	6	—	17
	32	—	—	—	4	7	3	14
	n_x	2	10	6	64	15	3	$n = 100$

6.	$y \backslash x$	10	12	14	16	18	20	n_y
	10	5	1	—	—	—	—	6
	30	—	6	2	—	—	—	8
	50	—	—	5	40	5	—	50
	70	—	—	2	8	7	—	17
	90	—	—	—	4	7	8	19
	n_x	5	7	9	52	19	8	$n = 100$

	$y \backslash x$	15	25	35	45	55	65	n_y
	15	4	2	—	—	—	—	6
	20	—	6	4	—	—	—	10
7.	25	—	—	6	45	2	—	53
	30	—	—	2	8	6	—	16
	35	—	—	—	4	7	4	15
	n_x	4	8	12	57	15	4	$n = 100$

	$y \backslash x$	5	10	15	20	25	30	n_y
	20	1	4	—	—	—	—	5
	30	—	7	3	—	—	—	10
8.	40	—	—	2	50	2	—	54
	50	—	—	1	10	6	—	17
	60	—	—	—	4	7	3	14
	n_x	1	11	6	64	15	3	$n = 100$

Задачи.

1. Себестоимость y (в $) одного экземпляра книги в зависимости от тиража x (в тыс. экземпляров) характеризуется статистикой, собранной издательством:

x_i	1	2	3	5	10	20	30	50	100	200
y_i	10.15	5.52	4.08	2.85	2.11	1.62	1.41	1.30	1.21	1.15

С помощью метода наименьших квадратов оценить коэффициенты θ_1, θ_2 гиперболической зависимости $y = \theta_1 + \theta_2/x$, дисперсию ошибок наблюдений σ^2, матрицу вариаций оценок и построить 90%-ные доверительные интервалы для θ_1, θ_2, а также для величины y, для различных x_i.

2. Сырье, поступившее на завод из карьера, содержит два полезных компонента — минералы A и B. Результаты анализа 10 партий сырья, поступившего в разное время из разных мест карьера представлены в таблице (ξ, η — соответственно %-ное содержание минералов A, B в образцах):

№ партии	1	2	3	4	5	6	7	8	9	10
ξ	67	54	72	64	39	22	58	43	46	34
η	24	15	23	19	16	11	20	16	17	13

Оценить коэффициент корреляции $cor(\xi, \eta)$ и функции регрессии η на ξ, и ξ на η.

3. В таблице указаны различные количества фосфора x_i, внесенные в почву девяти опытных участков, и количества фосфора в кукурузе y_i, выросшей на различных участках через 38 дней:

y_i	64	71	54	81	76	93	77	95	109
x_i	1	4	5	9	11	13	23	23	28

Предполагается, что модель зависимости y от x линейна:

$$y_i = \theta_1(x_i - \overline{x}) + \theta_2 + \varepsilon_i,\ i = \overline{1,9},\ \text{где}\ \overline{x} = \sum_{i=1}^{9} \frac{x_i}{9},$$

параметры θ_1, θ_2 — неизвестны, а случайные ошибки ε_i некоррелированы, имеют нулевые средние и одинаковые дисперсии σ^2. Оценить θ_1, θ_2, σ^2 по методу наименьших квадратов. Точки (x_i, y_i) и оценку $\widehat{\theta}_1(x_i - \overline{x}) + \widehat{\theta}_1$ прямой линии регрессии изобразить графически в системе координат xOy.

Приложения.

Законы распределения вероятностей, наиболее распространенные в практике статистических исследований.

№	Название закона	Общая схема формирования случайных данной природы	Примеры реальных признаков, подчиняющихся данному закону	Аналитическое задание закона
1.	Биномиальный $Bi(n,\theta)$	Число x появления интересующего нас события в последовательности из n независимых экспериментов, когда вероятность появления этого события в одном испытаннии остается постоянной и равна θ.	1. Число дефектных изделий в партии определенного объема, отобранной из массовой продукции. 2. Число объектов, обладающих заданной комбинацией свойств, оказавшихся среди n случайно отобранных.	$f(x,\theta) = C_n^x \theta^x (1-\theta)^{n-x}$, $x = \overline{0,n}$, $0 < \theta < 1$.
2.	Гипергеометрический $\nu_{M,N}(m)$	Число объектов $\nu_{M,N}(m)$, обладающих заданным свойством среди m объектов, случайно извлеченных из совокупности N объектов, M из которых обладают этим свойством.	Число дефектных изделий в выборке объема m, случайно извлеченной из партии изделий объема N, в которой оказалось M дефектных изделий.	$f(x) = \dfrac{C_M^x \cdot C_{N-M}^{m-x}}{C_N^M}$, $x = \max(0, m-(N-M), \ldots, \min(m, M))$.

3. Пуассона $\Pi(\theta)$	Число наступления, интересующего нас события за единицу времени, когда факт наступления такого события в данном эксперименте не зависит от того, сколько раз и в какие моменты времени оно наступило в прошлом и не влияет на будущее.	1. Число сбоев, отлаженного производственного процесса в единицу времени. 2. Число требований, поступающих в единицу времени в систему массового обслуживания. 3. Число несчастных случаев, происходящих в единицу времени в данной совокупности.	$f(x,\theta) = e^{-\theta}\,\dfrac{\theta^x}{x!}$, $x = 0, 1, 2, \ldots;$ $0 < \theta < \infty.$
4. Нормальный (гауссовский) $\mathcal{N}(\theta_1, \theta_2^2)$	Значения случайной величины формируются под действием большого числа взаимно независимых случайных факторов, причем сила воздействия каждого отдельного фактора мала и не может превалировать среди остальных, а характер воздействия — аддитивный.	1. Отклонение от номинала в значениях параметров изделий производимых в условиях стационарного массового производства. 2. Ошибки измерения. 3. Ошибки при стрельбе по цели.	$f(x,\theta_1,\theta_2) = \dfrac{1}{\theta_2\sqrt{2\pi}}\, e^{-\dfrac{(x-\theta_1)^2}{2\theta_2^2}}$, $-\infty < \theta_1 < \infty,$ $0 < \theta_2 < +\infty.$

5. Равномерный (прямоугольный) $R(\theta_1, \theta_2)$	Вероятность получить наблюдение из любой окрестности, заданного диапазона $[\theta_1, \theta_2]$ возможных значений данной случайной величины, зависит только от ширины этой окрестности и не зависит от того, в каком именно месте диапазона $[\theta_1, \theta_2]$ она располагается.	1. Ошибки округления при проведении числовых расчетов с фиксированным числом десятичных знаков. 2. Время ожидания обслуживания при точном периодическом включении обслуживающего устройства и при равномерно случайном прибытии заявки на обслуживание в этом интервале.	$f(x, \theta_1, \theta_2) = \frac{1}{\theta_2 - \theta_1}$, $-\infty < \theta_1 < \theta_2 < +\infty$.
6. Экспоненциальный (показательный) $\Gamma(\theta, 1)$	Случайный промежуток времени между двумя событиями пуассоновского типа (см. №3).	1. Время обслуживания в некоторых системах массового обслуживания. Долговечность изделия, работающего в нормальном режиме эксплуатации.	$f(x, \theta) = \begin{cases} e^{-\theta} \frac{\theta^x}{x!}, & x \geq 0, \\ 0, & x < 0, \end{cases}$ $0 < \theta < \infty$.

7. Коши $K(\theta)$		Отношение двух независимых нормальных случайных величин из $\mathcal{N}(0,1)$ подчиняются распределению Коши с $\theta = 0$, $c = 1$.	$f(x,\theta) = \frac{1}{\pi}\frac{c}{c^2 + (x-\theta)^2}$, $-\infty < \theta < +\infty$.
8. "Хи-квадрат" с m степенями свободы $\chi^2(m)$	$\chi^2(m) = \sum_{i=1}^{m} \xi_i^2$, $\xi_i \in \mathcal{N}(0,1), i = \overline{1,m}$.	1. Мера отклонения эмпирического распределения от теоретического. 2. Пронормированная выборочная дисперсия, построенная по выборке из нормальной совокупности.	$f(x) = \frac{2^{-\frac{m}{2}}}{\Gamma(\frac{m}{2})} x^{\frac{m}{2}-1} e^{-\frac{x}{2}}$, $0 \le x < \infty$.

9. Стьюдента с m степенями свободы $t(m)$	$t(m) = \dfrac{\xi_0}{\left(\frac{1}{m}\sum\limits_{i=1}^{m}\xi_i^2\right)^{1/2}},$ $\xi_i \in \mathcal{N}(0,1), i = \overline{0,m}.$	1. Нормированная мера расхождения двух выборочных средних, вычисленных по двум независимым выборкам из нормальной совокупности. 2. Нормированное отклонение выборочного среднего (построенного по выборке из нормальной генеральной совокупности) от соответствующего теоретического значения.	$f(x) = \dfrac{1}{\sqrt{\pi m}} \dfrac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \left(1 + \dfrac{x^2}{m}\right)^{-\frac{m+1}{2}},$ $-\infty < x < \infty.$
10. Фишера с числом степеней свободы m_1 и m_2 $F(m_1, m_2)$	$F(m_1,m_2) = \dfrac{\frac{1}{m_1}\sum\limits_{i=1}^{m_1}\xi_i^2}{\frac{1}{m_2}\sum\limits_{i=m_1+1}^{m_1+m_2}\xi_i^2},$ $\xi_i \in \mathcal{N}(0,1),$ $i = \overline{1, m_1+m_2}.$	Отношение двух выборочных дисперсий, построенным по двум независимым выборкам, извлеченным из нормальной генеральной совокупности.	$f(x) = \dfrac{\Gamma(\frac{m_1+m_2}{2})}{\Gamma(\frac{m_1}{2})\Gamma(\frac{m_2}{2})} \dfrac{m_1^{\frac{m_1}{2}} m_2^{\frac{m_2}{2}} x^{\frac{m_1}{2}-1}}{(m_1 x + m_2)^{\frac{m_1+m_2}{2}}},$ $0 \le x < \infty.$

Литература

+[1] Боровков А.А. *Теория вероятностей.* - М.: Наука., 1986. - 432 с.

[2] Гнеденко Б.В. *Курс теории вероятностей.* - М.: Наука, 1969. - 400 с.

[3] Климов Г.П. *Теория вероятностей и математическая статистика.* - М.: МГУ, 1983. -328 с.

[4] Розанов Ю.А. *Теория вероятностей, случайные процессы, математическая статистика.* - М.: Наука, 1985. - 320 с.

[5] Севастьянов Б.А. *Курс теории вероятностей и математической статистики.* - М.: Наука, 1982. - 256 с.

[6] Чистяков В.П. *Курс теории вероятностей.* - М.: Наука, 1987. - 240 с.

+[7] Ширяев А.Н. *Вероятность.* - М.: Наука, 1980.- 575 с.

[8] Феллер В. *Введение в теорию вероятностей и ее приложения.* - М.: Мир, 1984. - 527 с. (т.1)

[9] Мешалкин Л.Д. *Сборник задач по теории вероятностей.* - М.: МГУ, 1963.- 156 с.

[10] Прохоров А.В., Ушаков В.Г., Ушаков Н.Г. *Задачи по теории вероятностей.* - М.: Наука, 1986.- 327 с.

+[11] Зубков А.М., Севастьянов Б.А., Чистяков В.П. *Сборник задач по теории вероятностей.* - М.: Наука, 1989. - 318 с.

[12] Емельянов Г.В., Скитович В.П. *Задачник по теории вероятностей и математической статистике.* - Л.: ЛГУ, 1967. - 330 с.

[13] Ивченко Г.И., Медведев Ю.И. *Математическая статистика.* - М.: Высш.школа, 1984. - 248 с.

[14] Козлов М.В., Прохоров А.В. *Введение в математическую статистику.* - М.: МГУ, 1987. - 264 с.

[15] Харин Ю.С., Хацкевич Г.А., Лобач В.И. *Сборник задач по теории вероятностей, случайных процессов и математической статистике.* - Мн.: БГУ, 1995. - 99 с.