МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ

«БРАТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Кафедра управления в технических системах

УТВЕРЖДАЮ:
Проректор по учебной работе
_____________ Е.И. Луковникова
«31» мая 2019 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Б1.В.ДВ.05.01

НАПРАВЛЕНИЕ ПОДГОТОВКИ

27.03.04 Управление в технических системах

ПРОФИЛЬ ПОДГОТОВКИ

Управление и информатика в технических системах

Программа академического бакалавриата

Квалификация (степень) выпускника: бакалавр

Программа составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 27.03.04 Управление в технических системах от 20.10.2015 г № 1171 и учебным планом ФГБОУ ВО «БрГУ» от 01.04.2019 г № 196 для заочной формы обучения набора 2019 года

СОДЕРЖАНИЕ ПРОГРАММЫ **Стр.**

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Вид деятельности выпускника

Дисциплина охватывает круг вопросов, относящихся к научно- исследовательской и проектно- конструкторскому видам профессиональной деятельности выпускника в соответствии с компетенциями и видами деятельности, указанными в учебном плане.

Цель дисциплины

Формирование у обучающихся знаний о сборе, обработке и представлении экспериментальных данных.

Задачи дисциплины

Сформировать у обучающихся знания, умения, навыки использования основных приемов сбора, обработки и анализа данных, полученных в ходе профессиональной деятельности.

Код компетенции	**Содержание компетенций**	**Перечень планируемых результатов обучения по дисциплине**
1	**2**	**3**
ОПК-2	способность выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, привлекать для их решения соответствующий физико-математический аппарат	**Знать:** -Способность выявлять естественнонаучную сущность процесса, - Физико-математические закономерности процессов, -Способы оценки данных. **Уметь:** -Получать необходимую информацию, обрабатывать ее, - Использовать полученные знания на практике, -Проводить регрессионный, дисперсионный, многофакторный корреляционно-регрессионный анализ полученных данных. **Владеть:** - Достаточным уровнем понимания материала, и способностью выявлять сущность проблем, - Соответствующим физико-математическим аппаратом для решения практических задач.
ПК-2	способность проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления	**Знать:** - Основные приемы обработки данных, -Основные законы распределения вероятностей, -Методы обработки данных,

		-Способы представления экспериментальных данных. **Уметь:** - Проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления **Владеть:** -Способностью применять стандартные программные средства с целью получения математических моделей

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина Б1.В.ДВ.05.01 Математическая статистика относится к вариативной части.

Дисциплина Математическая статистика базируется на знаниях, полученных при изучении дисциплин: Математика, Математические модели и методы..

Основываясь на изучении перечисленных дисциплин, Математическая статистика представляет основу для изучения дисциплин: Автоматизация технологических процессов и производств, Метрология и измерительная техника.

Такое системное междисциплинарное изучение направлено на достижение требуемого ФГОС уровня подготовки по квалификации бакалавр.

3. РАСПРЕДЕЛЕНИЕ ОБЪЕМА ДИСЦИПЛИНЫ

3.1. Распределение объема дисциплины по формам обучения

Форма обучения	*Курс*	*Семестр*	*Трудоемкость дисциплины в часах*						*Контрольная работа*	*Вид промежуточной аттестации*
			Всего часов (с экз.)	*Аудиторных часов*	*Лекции*	*Лабораторные работы*	*Практические занятия*	*Самостоятельная работа*		
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Очная	3	5	180	68	17	34	17	112	-	Экзамен
Заочная	2	-	180	15	5	5	5	165	-	Экзамен
Заочная (ускоренное обучение)	1	-	180	11	5	3	3	169	-	Экзамен
Очно-заочная	-	-	-	-	-	-	-	-	-	-

3.2. Распределение объема дисциплины по видам учебных занятий и трудоемкости

Вид учебных занятий	*Трудо-емкость (час.)*	*в т.ч. в интерактивной, активной, иннова-циионной формах, (час.)*	*Распределение по семестрам, час*
			5
1	2	3	4
I. Контактная работа обучающихся с преподавателем (всего)	68	11	68
Лекции (Лк)	17	2	17
Лабораторные работы (ЛР)	34	7	34
Практические занятия (ПЗ)	17	2	17
Групповые (индивидуальные) консультации	+	-	+
II.Самостоятельная работа обучающихся (СР)	58	-	58
Подготовка к лабораторным работам	25	-	25
Подготовка к практическим занятиям	25	-	25
Подготовка к экзамену в течение семестра	8	-	8
III. Промежуточная аттестация экзамен	54	-	54
Общая трудоемкость дисциплины час.	180	-	180
зач. ед.	5	-	5

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Распределение разделов дисциплины по видам учебных занятий
- для очной формы обучения:

№ раз-дела и темы	*Наименование раздела и тема дисциплины*	*Трудоем-кость, (час.)*	*Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость; (час.)*			
			учебные занятия			*Самостоя тельная работа обучающихся*
			лекции	*лаборатор ные работы*	*практичес кие занятия*	
1	2	3	4	5	6	7
1.	**Основные сведения**	**16**	**2**	**4**	**2**	**8**
1.1.	Введение. Случайные величины.	3,5	0,5	-	-	3
1.2.	Числовые характеристики случайной величины.	6,5	0,5	4	1	1
1.3.	Статистическое распределение выборки.	6	1	-	1	4
2.	**Основные законы распределения вероятностей**	**23**	**4**	**6**	**3**	**13**
2.1.	Биноминальное распределение. Распределение Пуассона. Показательное распределение.	5	1	-	-	4
2.2.	Нормальное распределение. Распределение кси-квадрат.	10	2	6	-	2
2.3.	Распределение Стьюдента. F-распределение. Статистические оценки	8	1	-	3	4

3.	**Методы получения точечных оценок. Интервальные оценки**	**20**	**4**	**-**	**3**	**13**
3.1.	Метод максимального правдоподобия. Метод наименьших квадратов.	10	2	-	1	7
3.2.	Интервальные оценки.	10	2	-	2	6
4.	**Проверка статистических гипотез**	**18**	**2**	**-**	**3**	**13**
4.1.	Этапы проверки гипотез. Проверка гипотезы о равенстве выборочной средней.	10	1	-	3	6
4.2.	Проверка гипотезы о значении мат. ожидании. Проверка гипотезы о значении дисперсии.	8	1	-	-	7
5.	**Однофакторный, двухфакторный анализ**	**24**	**2**	**12**	**3**	**7**
5.1.	Виды зависимостей между признаками. Однофакторный дисперсионный анализ.	11	1	6	1	3
5.2.	Двухфакторный дисперсионный анализ.	13	1	6	2	4
6.	**Корреляционно- регрессионный анализ**	**25**	**3**	**12**	**3**	**7**
6.1.	Многофакторный корреляционно-регрессионный анализ. Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции	13,5	2,5	6	2	3
6.2.	Регрессионная модель. Построение регрессионной модели.	11,5	0,5	6	1	4
	ИТОГО	**126**	**17**	**34**	**17**	**58**

- для заочной формы обучения:

№ раз-дела и темы	***Наименование раздела и тема дисциплины***	***Трудоем-кость, (час.)***	***Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость; (час.)***			
			учебные занятия			***Самостоятельная работа обучающихся***
			лекции	***лабораторные работы***	***практические занятия***	
1	2	3	4	5	6	7
1.	**Основные сведения**	**26,5**	**0,5**	**0,5**	**0,5**	**25**
1.1.	Введение. Случайные величины.	8	-	-	-	8
1.2.	Числовые характеристики случайной величины.	9	0,5	0,5	-	8
1.3.	Статистическое распределение выборки.	9,5	-	-	0,5	9
2.	**Основные законы распределения вероятностей**	**26,5**	**0,5**	**0,5**	**0,5**	**25**
2.1.	Биноминальное распределение. Распределение Пуассона. Показательное распределение.	8	-	-	-	8
2.2.	Нормальное распределение. Распределение кси-квадрат.	9	0,5	0,5	-	8
2.3.	Распределение Стьюдента. F-распределение. Статистические оценки	9,5	-	-	0,5	9

3.	**Методы получения точечных оценок. Интервальные оценки**	**27**	**1**	**-**	**1**	**25**
3.1.	Метод максимального правдоподобия. Метод наименьших квадратов.	13	-	-	-	13
3.2.	Интервальные оценки.	14	1	-	1	12
4.	**Проверка статистических гипотез**	**27**	**1**	**-**	**1**	**25**
4.1.	Этапы проверки гипотез. Проверка гипотезы о равенстве выборочной средней.	13,5	0,5	-	1	12
4.2.	Проверка гипотезы о значении мат. ожидании. Проверка гипотезы о значении дисперсии.	13,5	0,5	-	-	13
5.	**Однофакторный, двухфакторный анализ**	**29**	**1**	**2**	**1**	**25**
5.1.	Виды зависимостей между признаками. Однофакторный дисперсионный анализ.	13,5	0,5	1	-	12
5.2.	Двухфакторный дисперсионный анализ.	15,5	0,5	1	1	13
6.	**Корреляционно- регрессионный анализ**	**35**	**1**	**2**	**1**	**31**
6.1.	Многофакторный корреляционно-регрессионный анализ. Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции	17	0,5	1	0,5	15
6.2.	Регрессионная модель. Построение регрессионной модели.	18	0,5	1	0,5	16
	ИТОГО	**171**	**5**	**5**	**5**	**156**

- для заочной формы (ускоренное обучение) обучения:

№ раз-дела и темы	***Наименование раздела и тема дисциплины***	***Трудоем-кость, (час.)***	***Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость; (час.)***			
			учебные занятия			***Самостоятельная работа обучающихся***
			лекции	***лабораторные работы***	***практические занятия***	
1	2	3	4	5	6	7
1.	**Основные сведения**	**25,5**	**0,5**	**-**	**-**	**25**
1.1.	Введение. Случайные величины.	8,5	0,5	-	-	8
1.2.	Числовые характеристики случайной величины.	8	-	-	-	8
1.3.	Статистическое распределение выборки.	9	-	-	-	9
2.	**Основные законы распределения вероятностей**	**26**	**1**	**-**	**-**	**25**
2.1.	Биноминальное распределение. Распределение Пуассона. Показательное распределение.	8	-	-	-	8
2.2.	Нормальное распределение. Распределение кси-квадрат.	9	1	-	-	8
2.3.	Распределение Стьюдента. F-распределение. Статистические оценки	9	-	-	-	9

3.	**Методы получения точечных оценок. Интервальные оценки**	**26**	**1**	**-**	**-**	**25**
3.1.	Метод максимального правдоподобия. Метод наименьших квадратов.	12,5	0,5	-	-	12
3.2.	Интервальные оценки.	13,5	0,5	-	-	13
4.	**Проверка статистических гипотез**	**26**	**1**	**-**	**-**	**25**
4.1.	Этапы проверки гипотез. Проверка гипотезы о равенстве выборочной средней.	13	1	-	-	12
4.2.	Проверка гипотезы о значении мат. ожидании. Проверка гипотезы о значении дисперсии.	13	-	-	-	13
5.	**Однофакторный, двухфакторный анализ**	**33**	**1**	**1**	**1**	**30**
5.1.	Виды зависимостей между признаками. Однофакторный дисперсионный анализ.	22	1	1	-	20
5.2.	Двухфакторный дисперсионный анализ.	11	-	-	1	10
6.	**Корреляционно- регрессионный анализ**	**34,5**	**0,5**	**2**	**2**	**30**
6.1.	Многофакторный корреляционно-регрессионный анализ. Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции	12,5	0,5	1	1	10
6.2.	Регрессионная модель. Построение регрессионной модели.	22	-	1	1	20
	ИТОГО	**171**	**5**	**3**	**3**	**160**

4.2. Содержание дисциплины, структурированное по разделам и темам

РАЗДЕЛ 1. Основные сведения

ВВЕДЕНИЕ

Математическая статистика – раздел математики, изучающий методы оценивания и сравнения распределений случайных величин и их характеристик по наблюдениям (выборкам этих случайных величин).

Статистика изучает как неизвестные параметры случайных явлений, так и достоверность априорных (сделанных заранее) предположений (гипотез) по выборкам – некоторым пробным испытаниям.

Источником информации для матстатистики служит статистическое наблюдение – первая стадия статистического исследования, представляющая собой обоснованный сбор информации об явлениях и процессах в изучаемом объекте. В результате статистического наблюдения должна быть получена объективная, сопоставимая и достаточно полная информация, позволяющая на последующих этапах исследования разработать обоснованные выводы о характере и закономерностях развития изучаемого явления.

Изучение статистических закономерностей – важнейшая задача статистики, которую она решает с помощью особых методов, видоизменяющихся в зависимости от характера исходной информации и целей познания. Знание характера и силы связей позволяет управлять различными техническими процессами и предсказывать их развитие.

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ.

Случайной величиной X называют величину, которая случайно принимает какое-то значение из совокупности своих значений, ее закон распределения может быть задан функцией распре[illegible]ия $F(x) = P(X < x)$. Функция распределения $F(x)$ - неубывающая, определенная на в[illegible]числовой оси, при этом $F(-\infty) = 0$ и $F(\infty) = 1$.

Дискретными называют случайные величины, значениями которых являются только отдельные точки числовой оси. Значениями непрерывной случайной величины могут быть любые точки какого-то интервала на числовой оси.

Закон распределения дискретной случайной величины X можно определить с помощью ряда распределения, заданного в виде таблицы 1:

Таблица 1.

X	x_1	x_2	x_3	…	x_n
P	p_1	p_2	p_3	…	p_n

В первой строке этой таблицы указаны все значения x_i дискретной случайной величины X, а во второй – вероятности p_i принятия случайной величиной соответствующих значений x_i. Сумма этих вероятностей равна единицы.

На основе ряда распределения можно получить функцию распределения дискретной случайной величины X:

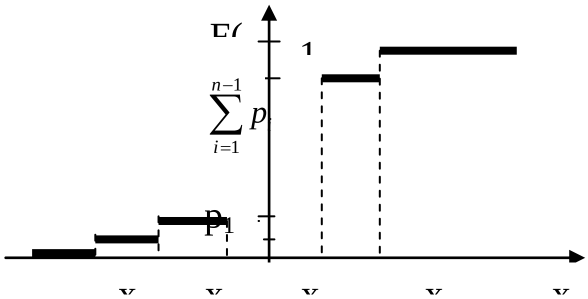
$$F(x) = \sum_{x_i \langle x} p_i$$

Данную формулу можно записать в следующем виде:

$$F(x) = \begin{cases} 0, & x \le x_1, \\ p_1, & x_1 \langle x \le x_2, \\ p_1 + p_2, & x_2 \langle x \le x_3, \\ \text{................................} & \\ \sum_{i=1}^{n-1} p_i, & x_{n-1} \langle x \le x_n, \\ 1, & x \rangle x_n. \end{cases}$$

График функции распределения дискретной случайной величины представляет собой ступенчатую линию, рис.1.

Рис.1.



Закон распределения непрерывной случайной величины задается или функцией распределения, или **функцией плотности вероятности.** Функция распределения непрерывной случайной величины X представляется в виде интеграла:

$$F(x) = \int_{-\infty}^{\infty} f(t)dt$$

где f(x)>0 – функция плотности вероятности. График этой функции всегда охватывает фигуру, площадь которой равна единице, т.к. $F(\infty) = 1$

Числовые характеристики случайной величины

1. Математическое ожидание – M(X) – серединное значение совокупности значений случайной величины (среднее значение) и для дискретных случайных величин вычисляется по формуле:

$$M(X) = \sum_{i=1}^{n} x_i p_i$$

где x_i – значение дискретной случайной величины, p_i – вероятность того, что случайная величина примет значение x_i, n – количество значений случайной величины.

Среднее значение непрерывной случайной величины определяется как:

$$M(X) = \int_{-\infty}^{\infty} x * f(x)dx$$

Математическое ожидание случайных величин имеет следующие свойства:

1) M(C)=C, C=const;
2) M(CX)=C*M(X) – постоянную можно выносить из под знака среднего значения;
3) M(X+У)=M(X)+M(У) – мат. ожидание суммы случайных величин равно сумме мат. Ожиданий случайных величин;
4) Если случайная величина $X \geq 0$, то и среднее значение $M(X) \geq 0$.

2. Дисперсия – D(X) – характеризует степень рассеяния значений случайной величины от своего среднего значения и для дискретных случайных величин определяется:

$$D(X) = M(X^2) - (M(X))^2.$$

Для непрерывных случайных величин дисперсия определяется как:

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - (M(X))^2$$

Дисперсия имеет следующие свойства:

1) Дисперсия константы равна нулю: D(X)=0;
2) Дисперсия всегда неотрицательна: $D(X) \geq 0$;
3) Постоянную можно вынести из под знака дисперсии, возведя ее в квадрат: $D(CX)=C^2*D(X)$;
4) Изменение случайной величины на постоянную не изменяет ее дисперсию: D(C+X)=D(X).

3. Среднее квадратичное отклонение – корень квадратный из дисперсии и означает абсолютное среднее отклонение случайной величины от своего среднего значения.

4. Мода – M_0 – значение случайной величины, которое встречается чаще всего, т.е. имеет максимальную вероятность (для дискретной случайной величины) или максимум функции плотности вероятности в данной точке (при непрерывной случайной величине).

5. Медиана – делит область значений случайной величины на две равные по вероятности части.

СТАТИСТИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ ВЫБОРКИ

Множество всевозможных значений случайной (дискретной или непрерывной) величины называется **генеральной совокупностью** данной величины.

Для изучения некоторого количественного признака X случайной величины из генеральной совокупности извлекают произвольным образом n его значений:

$$x_1, x_2, \ldots, x_n.$$

Это множество значений называют **выборкой** (выборочной совокупностью) объема n, сами значения x_i – вариантами, $1 \le i \le n.$ Разность между наибольшим и наименьшим значениями вариант называется **размахом** вариации. Для того, чтобы по выборке можно было достаточно уверенно судить об исследуемой случайной величине, выборка должна быть представительной, т.е. объекты выборки должны достаточно хорошо представлять генеральную совокупность. Достигается это тогда, когда все объекты генеральной совокупности имеют одинаковую вероятность попадания в выборку.

Для геометрического (наглядного) изображения статистического распределения выборки пользуются понятиями: **полигона** для дискретного распределения признака X случайной величины и **гистограммы** для интервального распределения X.

Полигоном частот выборки дискретного случая называют ломаную, отрезки которой соединяют точки (x_1,n_1), $(x_2,n_2),\ldots,(x_n,n_n)$, где x_i – варианта, n_i – соответствующая ей частота, $1 \le i \le n.$

Полигоном относительных частот выборки называют ломаную, отрезки которой соединяют точки $(x_1,\omega_1),(x_2,\omega_2),\ldots,(x_n,\omega_n)$, где x_i – варианта, ω_i – ее относительная частота, $1 \le i \le n.$

Гистограммой частот выборки называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы, а высоты равны n_i / h - плотности частот. Площадь частичного i – го прямоугольника равна $h * n_i / h = n_i$, а площадь всей гистограммы равна объему выборки n.

Гистограммой относительных частот выборки называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длины h, а высоты равны отношению ω_i / h (плотности относительной частоты). Площадь частичного i-го прямоугольника равна $h(\omega_i/h)=\omega_i$ – относительной частоте вариант, попавших в i-й интервал. Площадь гистограммы относительных частот равна сумме всех относительных частот, т.е. единице.

Последовательность вариант, записанные в возрастающем (или убывающем) порядке, называют **вариационным рядом.**

РАЗДЕЛ 2. ОСНОВНЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

Явления, происходящие в природе и обществе и носящие случайный характер, в большинстве случаев подчиняются определенным законам. Зная распределения вероятностей интересующих нас явлений, можно сформулировать выводы об этих явлениях.

1. Биноминальное распределение

Это распределение наиболее распространенное среди дискретных распределений, служит математической моделью многих явлений и возникает в тех случаях, когда интересующее нас явление X происходят k раз в серии из n независимых испытаний $(0 \le k \le n)$.

Если событие X имеет вероятность появления $p(0 < p < 1)$, то вероятность появления X k-раз в n испытаниях выражается формулой:

$$P(X = k | n, p) = C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}, q = 1-p.$$

2. Распределение Пуассона

Это распределение имеет место в тех случаях, когда в течение определенного времени может происходить случайное число каких-то событий (радиоактивных распадов, телефонных вызовов, отказов оборудования и т.д.).

Формула для вычисления вероятностей того, что случайная величина У примет целое неотрицательное значение k (k=0,1,2,…), имеет вид:

$$P(Y = k | \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0,1,2,...,$$

где λ - параметр распределения Пуассона, λ>0.

3. Показательное (экспоненциальное) распределение

Это распределение применяется в двух областях.

Первая область связана с продолжительностью безотказной работы устройств, вторая – с задачами массового обслуживания (интервалами между поломками оборудования, ремонтами).

Плотность показательного распределения с параметром Q имеет вид:

$$p(x,Q) = Qe^{-Qx}, x \geq 0.$$

Функция показательного распределения P(X<x) имеет вид:

$$F(x,Q) = \begin{cases} 1 - e^{-Qx}, x \geq 0; \\ 0, \quad x < 0. \end{cases}$$

Показательное распределение среди других выделяется отсутствием «памяти», т.е. изделие, прослужившее продолжительностью времени t, прослужит дополнительное время s с той же вероятностью, с какой это изделие прослужило бы времени s с самого начала

$$P(X \geq s + t | X \geq t) = P(X \geq s).$$

4. Нормальное распределение

Нормальное распределение относится к числу наиболее распространенных и важных, оно часто используется для приближенного описания многих случайных явлений, например, для случайного отступления фактического размера изделия от номинального, рассеяния снарядов при артиллерийской стрельбе, т.е. в ситуациях, когда на интересующий результат воздействует большое количество независимых случайных факторов, среди которых нет сильно выделяющихся.

Случайная величина У имеет нормальное распределение вероятностей, если ее плотность распределения задается формулой:

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-a)^2}{2\sigma^2}}, -\infty < x < +\infty.$$

Отметим, что $\varphi(x)$ стремится к нулю при $x \to -\infty$ и $x \to +\infty$. График функции $\varphi(x)$ симметричен относительно точки *a*. При этом в точке *a* функция $\varphi(x)$ достигает своего максимума, который равен 1/($\sqrt{2\pi\sigma}$) .

Парламетр *a* характеризует положение графика функции на числовой оси (параметр положения). Параметр $\sigma (\sigma > 0)$ характеризует степень сжатия или растяжения графика плотности (параметр масштаба).

Свойства:

1. Математическое ожидание и дисперсия случайной виличены *x*, распределённой как $\mathrm{N}(a, \sigma^2)$, равны

$$M(Y) = a, D(Y) = \sigma^2.$$

2. Медиана нормального распределения равна *a*, так как плотность распредиления симметрична относительно точки *x = a*.

Особую роль играет нормального распределеие с параметрами $a = 0$ и $\sigma = 1$ т.е. распределение N (0,1), которое часто называют *стандартным* нормальным распределением. Плотность стандартного нормального распределения есть

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Функцию распределения стандартного нормального распределения равна

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-Y^2/2} dy.$$

5. Распределение χ^2 (кси-квадрат)

Пусть случайные величины $\tau_1, \tau_2, \ldots, \tau_n$ – независимы , и каждая из них имеет стандартное нормальное распределение N(0,1). Случайная величина χ^2_n, определенная как:

$$\chi_n^2 = \tau_1^2 + \ldots + \tau_n^2,$$

имеет распределение кси-квадрат с n степенями свободы (параметр распределения).

6.Распределение Стьюдента

Пусть случайные величины $\tau_0, \tau_1, \ldots, \tau_n$ – независимы, и каждая из них имеет стандартное нормальное распределение N(0,1). Случайная величина

$$t_n = \frac{\tau_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\tau_i^2}}$$

имеет распределение Стьюдента с числом степеней свободы n, n=1,2…

7. F – распределение.

Пусть $\mu_1, \ldots, \mu_m$; $\tau_1, \ldots, \tau_n$ (где m,n – натуральные числа) обозначают независимые случайные величины, каждая из которых распределена по стандартному закону N(0,1). Случайная величина $F_{m,n}$, определенная как

$$F_{m,n} = \frac{\frac{1}{m}(\mu_1^2 + \ldots + \mu_m^2)}{\frac{1}{n}(\tau_1^2 + \ldots + \tau_n^2)},$$

имеет F – распределение с числами степеней свободы m и n. F – распределение называют еще распределением дисперсионного отношения.

Статистические оценки.

Статистики распределения, такие как: матожидание, дисперсия, мода, медиана, среднее квадратичное отклонение используются в качестве приближенных значений соответствующих числовых характеристик случайной величины X. Характеристика выборки Q^* , используемая в качестве приближенного значения, неизвестной генеральной характеристики Q, называется её точечной статистической оценкой.

Пусть Q – некоторая генеральная характеристика; x_1, x_2, … ,x_n - выборка, f - функция, позволяющая вычислить Q^*_n (аналог Q) для данной выборки.

Точечной оценкой неизвестного параметра Q количественного признака случайной величены является число (или точка на прямой)

$$Q_n^* = f(x_1, x_2, \ldots, x_n).$$

Оценка Q^*_n генеральной характеристики Q называется состоятельной, если для любого $\varepsilon > 0$ выполняет равенство:

$$\lim_{n\to\infty} P(\left|Q_n^* - Q\right| < \varepsilon) = 1.$$

Несмещенной называют точечную оценку Q^*_n, математический ожидание которой равно оцениваемому параметру Q при любом объём выборки:$M(Q^*_n)=Q$.

Несмещённая оценка Q^*_n характеристики Q называется эффективной, если она среди всех несмещённых оценок характеристики Q обладает наименьшей дисперсией.

Смещенной называют точечную оценку Q^*_n, математическое ожидание которой не равно оцениваемому параметру: $M(Q^*_n) \neq Q$.

Характеристика Q^* выборки, используемая в качестве приближенного значения неизвестной характеристики Q для ГС, называется точечной статистической оценкой (точечной оценкой) Q,

$$Q^* = f(x_1,...,x_n).$$

Несмещенной оценкой генеральной средней (математического ожидания) служит выборочная средняя $\bar{x}_B = \dfrac{\sum_{i=1}^{k} n_i x_i}{n}$,

где x_i - варианта выборки; n_i - частота варианты x_i; $n = \sum_{i=1}^{k} n_i$ - объём выборки.

РАЗДЕЛ 3. МЕТОДЫ ПОЛУЧЕНИЯ ТОЧЕЧНЫХ ОЦЕНОК. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ.

Метод максимального правдоподобия. (ММП)

Пусть известен закон распределения случайной величины, однако неизвестны один или несколько параметров, входящих в уравнение этого закона. Используя данный метод можно найти оценку этих параметров.

1) Пусть случайная величина распределена по закону Пуассона:

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0,1,2,...,$$

где $x_1, x_2,\ldots, x_n$ – некоторые независимые значения X. При каком λ (λ - неизвестный параметр) максимальна вероятность того, что случайная величина X за n наблюдений примет именно эти значения $x_1, x_2,\ldots, x_n$?

Решение: Построим функцию правдоподобия:

$$L(x_1, x_2,..., x_n) = P((X = x_1);(X = x_2);...;(X = x_n)).$$

чем больше значение L $(x_1, x_2,\ldots, x_n)$, тем правдоподобнее (более вероятно) появление чисел $x_1, x_2,\ldots, x_n$ в результате n наблюдений.

Так как $x_1, x_2,\ldots, x_n$ - независимые значения X то,

$$P((X = x_1);(X = x_2),...,(X = x_n)) = P(X = x_1) \cdot P(X = x_2) \cdot ... \cdot P(X = x_n).$$

тогда:

$$L(x_1, x_2,..., x_n) = \frac{\lambda^{x1}}{x_1!} e^{-\lambda} \cdot ... \cdot \frac{\lambda^{xn}}{x_n!} e^{-\lambda} = \frac{\lambda^{x1+...+xn} \cdot e^{-n\lambda}}{x_1! \cdot ... \cdot x_n!}.$$

Функция L является функцией и параметра λ,

$$L(x_1, x_2,..., x_n, \lambda) = \frac{\lambda^{x1+...+xn} \cdot e^{-n\lambda}}{x_1! \cdot ... \cdot x_n!}.$$

Через $\hat{\lambda}$ обозначим искомое значение λ, т.е.

$$L(x_1, x_2,..., x_n, \hat{\lambda}) = \max_{\lambda} L(x_1, x_2,..., x_n, \lambda).$$

Логарифмируя и применяя производную, найдем:

$$\hat{\lambda} = \frac{x_1 + x_2 + ... + x_n}{n} = \bar{x}_B$$ - среднее арифметическое данной выборки.

2) пусть случайная величина X распределена по показательному закону, т.е. функция плотности равна:

$$f_x = \mu e^{-\mu x}, x > 0;$$

$x_1, x_2,\ldots, x_n$- некоторые независимые значения X. При каком μ (μ - неизвестный параметр) максимальна вероятность того, что случайная величина X за n наблюдений примет именно эти значения $x_1, x_2,\ldots, x_n$?

Решение: Функция правдоподобия в этом случае имеет вид:

$$L(x_1, x_2,..., x_n) = f_x(x_1) \cdot f_x(x_2) \cdot ... \cdot f_x(x_n) = \mu^n e^{-\mu(x1+x2+...+xn)}$$

Функция L зависит и от параметра μ, поэтому:

$$L(x_1, x_2, ..., x_n, \mu) = f_x(x_1) \cdot f_x(x_2) \cdot ... \cdot f_x(x_n) = \mu^n e^{-\mu(x1+x2+...+xn)}$$

Логарифмируя и применяя производную, получим искомое значение $\hat{\mu}$:

$$\hat{\mu} = \frac{n}{x_1 + x_2 + ... + x_n} = \frac{1}{\overline{x_B}}.$$

Метод наименьших квадратов

Пусть случайная величина X имеет нормальный закон распределения с параметрами a=M(X) и σ=$\sqrt{D(X)}$, т.е. функция плотности имеет вид:

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-a)^2}{2\sigma^2}},$$

причем a – неизвестно, a σ - известна. Пусть x_1, x_2,…, x_n – неизвестные значения случайной величины X. При каком a максимальна вероятность того, что случайная величина X за n наблюдений примет именно эти значения
x_1, x_2,…, x_n?

Решение: Функция правдоподобия имеет вид:

$$L(x_1, x_2, ..., x_n, a) = f_x(x_1) \cdot f_x(x_2) \cdot ... \cdot f_x(x_n) =$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\frac{1}{\sigma}\right)^n e^{\frac{(x1-a)^2+(x2-a)^2+...+(xn-a)^2}{2\sigma^2}}$$

Логарифмируя, получим:

$$\ln L = n \ln \frac{1}{\sqrt{2\pi}} + n \ln \frac{1}{\sigma} - \frac{1}{2\sigma^2} (\sum_{i=1}^{n} (x_i - a)^2)$$

Функция L достигает максимума в той же точке, что и $\ln L$, а $\ln L$, - максимума тогда и только тогда, когда $F(a) = \sum_{i=1}^{n} (x_i - a)^2 \to \min$.

Из-за последнего соотношения метод нахождения оценки «a» называется методом наименьших квадратов.

Точку минимума функции y=F(a) находят через производную:

$$\frac{dF(a)}{(da)^2} = 2n > 0 \quad при \quad a = \hat{a}.$$

отсюда следует, что $\hat{a} = \overline{x_B}$ - точка минимума для F(a), а следовательно, точка максимума для L, т.е.:

$$\hat{a} = \overline{x_B} = \frac{x_1 + ... + x_n}{n}.$$

Интервальные оценки

В матстатистике, чтобы получить интервальную оценку по сделанной выборке находится точечная оценка Q* неизвестной характеристики Q, затем задаются вероятностью γ и по определенным правилам находят такое число δ>0, чтобы имело место равенство:

$$P(|Q^* - Q| < \delta) = \gamma,$$

$$Q^* - \delta < Q < Q^* + \delta.$$

Число δ называется точностью оценки Q* (чем меньше δ, тем выше точность оценки); числа Q_1 и Q_2 называются доверительными границами; интервал (Q_1,Q_2) доверительным интервалом или доверительной оценкой характеристики Q, вероятность γ - доверительной вероятностью или надежностью интервальной оценки.

Таким образом, интервальной называют оценку, которая определяется двумя числами – концами интервала, покрывающего оцениваемый параметр.

Доверительным называют интервал, который с заданной надежностью γ покрывает заданный параметр.

1.Интервальной оценкой с надежностью γ математического ожидания, а нормально распределенного количественного признака X по выборочной средней $\overline{\chi}_B$ при известном среднем квадратическом отклонении σ генеральной совокупно служит доверительный интервал

$\overline{\chi}_{B}$ - t∘($\sigma/\sqrt{n}$)> a > $\overline{\chi}_{B}$ + t∘($\sigma/\sqrt{n}$),

где t∘($\sigma/\sqrt{n}$)= δ - точность оценки, n – объём выборки, t – значение аргумента функции Лапласа Φ (t) (см. приложение 2), при котором Φ (t) = γ/2;

2. Интервальной оценкой (с надежностью γ) среднего квадратического отклонения σ нормального распределенного количественного признака X по «исправленному» выборочному среднему квадратическому отклонению s служит доверительный интервал

s (1 – q) < σ < s (1+q) (при q < 1),

0 < σ < s (1 +q) (при q > 1).

где q находят по таблицам по заданным n и γ .

РАЗДЕЛ 4. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ.

Статистической гипотезой называют любое высказывание о виде неизвестного распределения или о параметрах известных распределений генеральной совокупности.

Гипотезы проверяются по выборке. **Нулевой (основной)** называют выдвинутую гипотезу H_0. **Конкурирующей (альтернативной)** называют гипотезу H_1, которая противоречит нулевой. **Простой** называют гипотезу, содержащую только одно предположение. **Сложной** называют гипотезу, которая состоит из конечного или бесконечного числа простых гипотез.

В итоге проверки гипотез могут быть допущены ошибки двух родов.

Ошибка первого рода состоит в том, что будет отвергнута правильная нулевая гипотеза. Вероятность ошибки первого рода обозначают через α и называют условием значимости.

Ошибка второго рода состоит в том, что будет принята неправильная нулевая гипотеза. Вероятность ошибки второго рода обозначают через β.

Статистическим критерием называют случайную величину К, которая служит для проверки гипотезы.

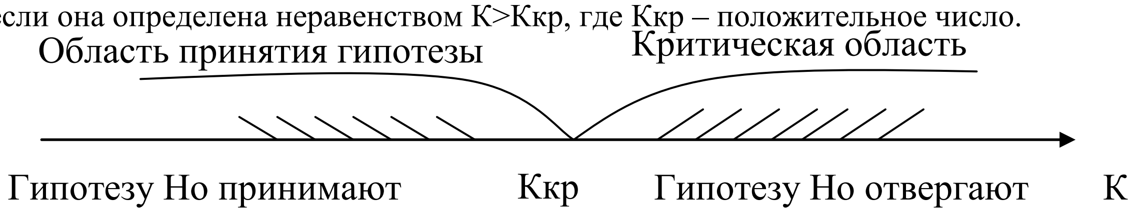
Наблюдаемым или эмпирическим значением Кнабл называют то значение критерия, которое вычислено по выборкам.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называют совокупность значений критерия, при которых нулевую гипотезу принимают.

Основной принцип проверки статистических гипотез: если наблюдаемое значение критерия принадлежит критической области, то нулевую гипотезу отвергают; если наблюдаемое значение критерия принадлежит области принятия гипотезы, то гипотезу принимают.

Критическими точками (границами) Ккр называют точки, отделяющие критическую область от области принятия гипотезы. Критическую область называют **правосторонней,** если она определена неравенством К>Ккр, где Ккр – положительное число.

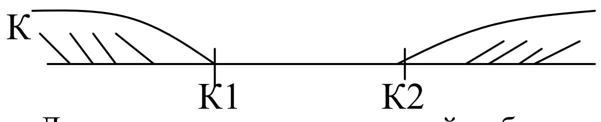


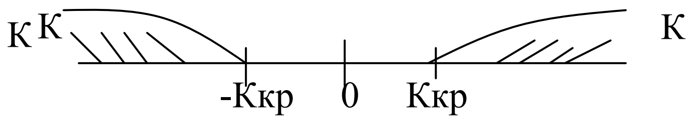
Критическую область называют **левосторонней**, если она определена неравенством К<Ккр, где Ккр отрицательное число.

Двусторонней называют критическую область, определяемую неравенствами К<К1, К>К2, где К1>К2. В частности, если критические точки симметричны относительно нуля, то двусторонняя критическая область определяется неравенствами, в предположении, что Ккр>0, как:

К<-Ккр, К>Ккр,

Или равносильным неравенством $|K| > Ккр$.





Для отыскания критической области задаются уровнем значимости α и ищут критические точки, исходя из следующих соотношений:

а) для правосторонней критической области:

$$P(K > Ккр) = \alpha, \quad (Ккр > 0);$$

б) для левосторонней критической области:

$$P(K < Ккр) = \alpha, \quad (Ккр < 0);$$

в) для двусторонней симметричной области:

$$P(K > Ккр) = \alpha/2, \quad (Ккр > 0);\ P(K < -Ккр) = \alpha/2.$$

Мощность критерия называют вероятность попадания критерия в критическую область при условии, что справедлива конкурирующая гипотеза. Другими словами, мощность критерия 1-β есть вероятность того, что нулевая гипотеза будет отвергнута, кроме того, верна конкурирующая гипотеза.

Этапы проверки гипотезы.

1 этап: Располагая выборочными данными $x_1, x_2, ..., x_n$ и руководствуясь условиями и целями задачи, формулируют две гипотезы: основную Но и альтернативную Н1

2 этап: Задаются уровнем значимости α, т.е. вероятностью того, что будет отвергнута правильная гипотеза Но (задается заранее и является малым числом: 0.05; 0.01; и т.д.). например $\alpha=0.05= Рн_0$ (Н1) означает, что из 100 случаев только в 5 случаях будет принята неправильная гипотеза Н1, а правильная Но отвергнута.

3 этап: По выборке находят критерий К=Кнабл, а по таблицам – Ккр и критическую область.

4 этап: Вывод: если Кнабл принадлежит критической области, то Но отвергают, если же Кнабл принадлежит области принятия, то Но принимают.

Проверка гипотезы о равенстве выборочной средней и гипотетической средней генеральной совокупности с нормальным распределением (дисперсия генеральной совокупности известно)

Правило1. Для того чтобы при заданном уровне значимости проверить нулевую гипотезу Но:$\alpha=\alpha_0$ - о равенстве генеральной средней α нормальной совокупности с известной дисперсией σ^2 гипотетическому (предполагаемому) значению α_0 при конкурирующей гипотезе Н1: $\alpha\neq\alpha_0$, надо: 1) вычислить наблюдаемое значение критерия

$$U_{набл} = \frac{(\bar{x} - \alpha_0)\sqrt{n}}{\sigma}$$

2) по таблице приложения функции Лапласа найти критическую точку $u_{кр}$ двусторонней критической области из равенства

$$\Phi(u_{кр}) = (1-\alpha)/2$$

3) если $|U_{набл}| < u_{кр}$ – нет оснований отвергнуть нулевую гипотезу; если $|U_{набл}| > u_{кр}$ – нулевую гипотезу отвергают.

Правило 2. При конкурирующей гипотезе Н1: $\alpha>\alpha_0$ критическую точку правосторонней критической области находят из равенства

$$\Phi(u_{кр})=(1-2\alpha)/2$$

Если $|U_{набл}| < u_{кр}$ – нет основания отвергнуть нулевую гипотезу. Если $|U_{набл}| > u_{кр}$ – нулевую гипотезу отвергают.

Правило 3. При конкурирующей гипотезе H1: $\alpha<\alpha_0$ сначала находят вспомогательную критическую точку $u_{кр}$ по правилу 2, а затем полагают границу левосторонней критической области

$$u'_{кр} = -u_{кр}$$

Если $|U_{набл}| > -u_{кр}$ – нет оснований отвергнуть нулевую гипотезу. Если $|U_{набл}| < -u_{кр}$ – нулевую гипотезу отвергают.

Проверка гипотезы о числовом значении математического ожидания нормального распределения при неизвестной дисперсии (критерий Стьюдента)

Пусть $x_1, x_2, ..., x_n$ - выборка, взятая из нормально распределенной генеральной совокупности с неизвестными параметрами а и σ; α - уровень значимости,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s^2 = \left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)\cdot\frac{1}{n-1},$$

α_0 - заданное число. Начальная гипотеза Ho: $\bar{x} = \alpha_0$, т.е. сравниваются числа $\bar{x}$ и α_0.

Правила проверки гипотезы Ho: $\bar{x} = \alpha_0$ зависят от вида альтернативной гипотезы H1.

Правило 1. Пусть H1: $\bar{x} \neq \alpha_0$. Тогда находим:

1) двустороннюю критическую область $(-\infty,-t_\gamma)$ и (t_γ, ∞) - критическое число t_γ определяется по таблице для γ и k, где γ=1-α, k=n-1;

2) $t_{набл} = \frac{\bar{x} - \alpha_0}{s}\sqrt{n}.$

3) Если $|t_{набл}| > t_\gamma$, , то Ho отвергаем. Если $|t_{набл}| < t_\gamma$, , то Ho не отвергаем.

Правило 2. Пусть H1: $\bar{x} > \alpha_0$. Тогда находим:

1) правостороннюю критическую область (t_γ,∞), t_γ определяется по таблице критерия Стьюдента для γ и k, где γ=1-2α, k=n-1;
2) $t_{набл}$ находится также как по правилу 1.
3). Если $t_{набл}>t_\gamma$, то Ho отвергаем. Если $t_{набл}<t_\gamma$, то Ho принимаем.

Правило 3. Пусть H1: $\bar{x} < \alpha_0$. Тогда находим:

1) критическую левостороннюю область $(-\infty, -t_\gamma)$, где t_γ вычисляется так же, как и в правилах 1 и 2.

2) $t_{набл}$ так же, как и в правилах 1 и 2.
3) Если $t_{набл}<-t_\gamma$, то Ho отвергается. Если $t_{набл}>-t_\gamma$,, то Ho принимается.

Проверка гипотезы о числовом значении дисперсии нормального распределения (критерий Пирсона или критерий χ^2)

Пусть случайная величина имеет нормальное распределение X=N(a,σ) с неизвестной дисперсией σ^2; $x_1,\ldots,x_n$ - результаты независимых наблюдений, проводимых в одинаковых условиях;

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Проверим гипотезу Ho: $\sigma^2 = \sigma_0^2$,

Где σ_0^2- некоторое заданное число.

В основе проверки этой гипотезы лежит сравнение σ_0^2 и s^2.

Правило 1. Пусть H1: $\sigma^2 \neq \sigma_0^2$. Тогда находим:

1) двустороннюю критическую область $\left(0, \chi_1^2\right), \left(\chi_2^2, \infty\right)$, где χ_1^2 определяется по таблице критерия Пирсона для $\gamma_1 = 1 - \alpha/2, \quad k = n-1;$ χ_2^2 находится по той же таблице для $\gamma_1 = \alpha/2, \quad k = n-1;$

2) $\chi^2_{набл} = \dfrac{(n-1)s^2}{\sigma_0^2}$ *вычисляетсяповыборке.*

3). Если $\chi^2_{набл}$ попадает в критическую область, то Но отвергаем, если нет, то Но принимаем.

Правило 2. Пусть альтернативная гипотеза имеет вид Н1: $\sigma^2 > \sigma_0^2$. Тогда находим:

1) правостороннюю критическую область (χ^2_γ, ∞), где критическое число χ^2_γ вычисляется по таблице для $\gamma = \alpha \quad и \quad k = n-1;$

2) Вычисляется так же как по правилу 1.

3). Если $\chi^2_{набл} > \chi^2_\gamma$, то Но отвергаем, если же $\chi^2_{набл} < \chi^2_\gamma$, то Но принимаем.

Правило 3. Пусть теперь альтернативная гипотеза имеет вид: Н1: $\sigma^2 < \sigma_0^2$. Тогда находим:

1) левостороннюю критическую область $\left(0, \chi^2_\gamma\right)$, где χ^2_γ вычисляется по таблице для $\gamma = 1 - \alpha \quad и \quad k = n-1;$

2) $\chi^2_{набл} = \dfrac{(n-1)s^2}{\sigma_0^2}$ *вычисляетсяповыборке.*

3). Если $\chi^2_{набл}$ принадлежит критической области, то принимается Н1. Если же нет, то принимается Но.

РАЗДЕЛ 5. ОДНОФАКТОРНЫЙ АНАЛИЗ. ДВУХФАКТОРЫЙ АНАЛИЗ.

ОДНОФАКТОРНЫЙ АНАЛИЗ

При исследовании зависимостей одной из наиболее простых является ситуация, когда можно указать только один фактор, влияющий на конечный результат, и этот фактор может принимать лишь конечное число значений (уровней). Типичный пример однофакторного анализа – сравнение по достигаемым результатам нескольких различных способов действия, направленных на достижение одной цели, н-р: несколько школьных учебников, несколько лекарств.

Для описания задач однофакторного анализа используется следующая терминология :

=то, что должно оказывать влияние на конечный результат называют фактором;

=конкретную реализацию фактора (н-р: определенный учебник) называют уровнем фактора или способом обработки;

=значения измеряемого признака называют откликом, а саму переменную – результативным признаком.

Для сравнения влияния факторов на результат необходим определенный статистический материал. Обычно его получают следующим образом: каждый из k способов обработки (наблюдений) применяют несколько раз (не обязательно одно и тоже число раз) к исследуемому объекту и регистрируют результаты. итогом подобных испытаний является k выборок.

Одной из главных целей однофакторного анализа является оценка величины влияния различных уровней фактора. Сначала проверяют, есть ли влияние вообще, нельзя ли расхождение результатов наблюдений объяснить действием чистой случайности. В качестве нулевой гипотезы здесь берут Но: данные столбцов принадлежат одному и тому же распределению.

ВИДЫ ЗАВИСИМОСТЕЙ МЕЖДУ ПРИЗНАКАМИ

Связи между явлениями и их признаками классифицируются по степени тесноты связи, направлению и аналитическому выражению.

Следует отметить два типа связей: функциональную и стохастическую.

Связь признака У с признаком Х называется функциональной, если каждому возможному значению независимого признака Х соответствует единственное значение зависимого признака У.

Функциональную связь от одного признака можно представить уравнением: $y_i = f(x_i)$.

Где y_i – значение результативного признака, f – известная функция, x_i – значение признака Х (i=1,2,…,n).

Зависимость величины У от Х называется стохастической (вероятностной), если каждому значению x независимой переменной Х соответствует не одно, а множество значений переменной У, причем заранее неизвестно какое из них переменная У примет.

Пусть при $X=x_i$ значениями У будут $y_{11}, y_{12},\dots,y_{1n1}$ и т.д.

Среднее множества значений $y_{11}, y_{12},\dots,y_{1n1}$ называют групповым средним переменной У при $X=x_i$ (это среднее значение первого столбца). Указанное среднее является условным математическим ожиданием случайной величины У при $X=x_i$:

$$\bar{y}_x = M(Y \mid X = x) = \frac{y_1 + y_2 + ... + y_{nx}}{n_x}$$

Аналогично определяется групповое среднее для значений Х, равных $x_2,\dots,x_k$.

Если при изменении x изменяются средние значения $\bar{y}_x$, то говорят, что имеет место корреляционная зависимость величины У от Х, а если же нет, то корреляционная связь между У и Х отсутствует.

Функция, φ(x)=M(Y|X=x) описывающая изменение среднего группового значения переменной У при изменении значений x переменной Х, называется функцией регрессии У на Х.

Модель стохастической связи определяется уравнением

$$\hat{y}_i = f(x_i) + \varepsilon_i \quad 1 \le i \le n,$$

Где $\hat{y}_i$ - расчетное значение результативного признака Y, $f(x_i)$ - часть результативного признака, сформировавшаяся под воздействием известных факторных признаков (одного или множества), ε_i- часть результативного признака, возникшая под воздействием неучтенных случайных факторов, а также неточности измерения.

Проявление стохастических связей подвержено действию закона больших чисел: лишь в большом числе испытаний случайности погасятся и зависимость, если она имеет существенную силу, проявится достаточно отчетлива.

ПРЯМЫЕ И ОБРАТНЫЕ СВЯЗИ

В зависимости от направления действия функциональные и стохастические связи могут быть прямыми и обратными. При прямой связи направление изменения результативного признака совпадает с направлением изменения признака –фактора, т.е. с увеличением факторного признака увеличивается и результативный, и наоборот, с уменьшением факторного признака уменьшается результативный.

ПРЯМОЛИНЕЙНЫЕ И КРИВОЛИНЕЙНЫЕ СВЯЗИ

По аналитическому выражению связи могут быть прямолинейными и криволинейными. При прямолинейной связи с возрастанием значения факторного признака происходит возрастание или убывание значений результативного признака. Математически такая связь представляется уравнением первой степени, а графически – прямой линией, по другому – линейная связь.

При криволинейных связях с возрастанием значения факторного признака возрастание или убывание результативного признака происходит неравномерно или же направление его изменения меняется на обратное. Геометрически такие связи представляются кривыми линиями.

ОДНОФАКТОРНЫЕ И МНОГОФАКТОРНЫЕ СВЯЗИ

По количеству факторов, действующих на результативный признак, связи различаются на однофакторные и многофакторные. Однофакторные (простые) связи обычно называются парными, т.к. рассматривается пара признаков – один влияющий фактор и один результативный.

В случае многофакторной (множественной) связи имеют ввиду, что все факторы действуют комплексно, т.е. одновременно и во взаимосвязи.

ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Однофакторный дисперсионный анализ позволяет по выборочным данным выяснить, влияет ли контролируемый фактор на результативный признак, и при наличии такого влияния оценить его степень с помощью так называемого коэффициента детерминации

$$R^2 = \frac{S^2_{факт}}{S^2}$$

показывающего какая доля общей дисперсии объясняется зависимостью результирующей переменной от фактора, где $S^2_{факт}$ – факторная дисперсия; S^2 – дисперсия результирующего признака.

Дисперсионный анализ используется для выявления влияния на изучаемый показатель некоторых факторов, обычно не поддающихся количественному измерению. Суть метода состоит в разложении общей вариации изучаемого показателя на части, соответствующие раздельному и совместному влиянию факторов, и статистическом изучении этих частей с целью выяснения приемлемости гипотез о существовании этих влияний. Модели дисперсионного анализа в зависимости от числа факторов классифицируются на однофакторные, двухфакторные и т.д.

По цели исследования модели делят на: детерминированные – здесь уровни всех факторов заранее фиксированы и проверяют именно их влияние; случайные – уровни каждого фактора получены как случайная выборка из генеральной совокупности уровней фактора, и смешанные – уровни одних факторов заранее фиксированы, а уровни других – случайная выборка.

На практике дисперсионный анализ применяют для установления влияния некоторого качественного фактора F, имеющего p уровней, на изучаемую величину X. Если выясняется урожайность зерновых культур, то фактором F является удобрение, а уровнем F – виды удобрений.

Основная идея дисперсионного анализа состоит в сравнении факторной дисперсии, порождаемой воздействием фактора, и остаточной дисперсией, обусловленной случайными причинами. Если различие между этими дисперсиями значимо, то фактор оказывает существенное влияние на X, в этом случае среднее наблюдаемых значений на каждом уровне различаются также значимо.

Сравнение нескольких дисперсий по выборкам, взятым из нормально распределенной генеральной совокупности

Пусть на количественный нормально распределенный признак X воздействует фактор F, имеющий p уровней $F_1, F_2, \ldots, F_p$. На каждом уровне Fi произведено n_i испытаний, 1<=i<=p, числа $n_1, \ldots, n_p$ не обязательно равные.

Требуется на уровне значимости α проверить основную гипотезу Ho об однородности дисперсий, т.е. гипотезу Ho: D(X1)=D(X2)=…=D(Xp), где - Xi - выборка соответствующая уровню Fi (i-й столбец, 1<=i<=p).

Алгоритм проверки следующий:

1. по групповым выборкам найдем исправленные выборочные дисперсии $s_1^2, s_2^2, \ldots, s_p^2$, где $s_i^2 = \frac{\hat{\sigma}_i^2 n_i}{n_i - 1}$, $\quad \hat{\sigma}_i^2 = \frac{(x_{1i} - \bar{x}_i)^2 + (x_{2i} - \bar{x}_i)^2 + \ldots + (x_{ni} - \bar{x}_i)^2}{n_i}$

$\hat{\sigma}_i^2$ - уровневая групповая дисперсия, i=1,2,…,p .

2. Находим

$$\overline{s^2} = \frac{k_1 s_1^2 + k_2 s_2^2 + ... + k_p s_p^2}{k} = \frac{\sum_{i=1}^{p} k_i s_i^2}{k}$$

где: $k_i=n_i-1$, $k=k_1+k_2+…+k_p$

Число $\overline{s^2}$ – средняя арифметическая исправленных дисперсий групп.

3. Вычислим:

$$V = 2.303\left(k \lg \overline{s^2} - \sum_{i=1}^{p} k_i \lg s_i^2\right),$$

$$C = 1 + \frac{1}{3(p-1)}\left(\sum_{i=1}^{p} \frac{1}{k_i} - \frac{1}{k}\right),$$

$$Б_{набл} = \frac{V}{C}$$

Случайная величина $Б_{набл}$ называется критерием Барлетта. Эта величина при условии справедливости Но аппроксимируется распределением χ^2 с p-1 степенями свободы, если для любого i, 1<=i<=p, n_i>=4.

Правило принятия решения:

На уровне значимости Но отклоняется, если $Б_{набл} > \chi^2(\alpha, p-1)$, $\chi^2_{кр}(\alpha, p-1)$– критическая точка правосторонней критической области. Если $Б_{набл} < \chi^2(\alpha, p-1)$,, то нет оснований отклонять Но.

ДВУХФАКТОРНЫЙ АНАЛИЗ

Двухфакторный анализ проводят для задач, в которых на результат влияют два фактора, каждый из которых принимает конечное число значений или уровней. Анализ выполняется для численной оценки этого влияния и выбора оптимального уровня фактора.

Пусть фактор А принимает k уровней, а фактор В n уровней (значений). Простейшая таблица двухфакторного анализа имеет вид:

Таблица 1.

	Обработки			
Блоки	1	2	…	k
1	x_{11}	x_{12}	…	x_{1k}
2	x_{21}	x_{22}	…	x_{2k}
…	…	…	…	…
n	x_{n1}	x_{n1}	…	x_{nk}

Приведенная таблица содержит nk наблюдений (в каждой клетке результат одного наблюдения). На практике часто рассматривают более сложные таблицы с повторными измерениями (в каждой клетке могут содержаться результаты нескольких наблюдений и обычно их заменяют средними арифметическими).

Каждое наблюдение x_{ij} представляется в виде:

$$x_{ij} = b_i + t_j + e_{ij}, \quad i = 1,..., n; \quad j = 1,..., k \qquad (1)$$

Где $b_1,…,b_n$ – результат влияния фактора В; $t_1,…,t_k$- результат влияния фактора А; e_{ij} - случайные величины, они одинаково распределены и независимы в совокупности, отражающие внутренне присущую наблюдениям изменчивость.

Равенство (1) можно представить в следующем виде:

$$x_{ij} = \mu + \beta_i + \tau_j + e_{ij}, \quad i = 1,.., n; \quad j = 1,.., k \qquad (2)$$

Где μ – среднее всей совокупности x_{ij}; β_i- отклонения результатов действия фактора В от μ; τ_j - отклонения результатов действия фактора А от μ.

Гипотеза Но: $\tau_1 = ... = \tau_k = 0$, т.е. влияние фактора А отсутствует.

РАЗДЕЛ 6. КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ

Корреляционный анализ и регрессионный анализ являются смежными разделами математики и предназначены для изучения зависимости величин по их выборочным данным.

В общем виде задача статистики в области изучения взаимосвязей состоит не только в количественной оценке их наличия и направления, то и в нахождении формулы влияния факторных признаков на результативный. Для ее решения применяют методы корреляционного и регрессионного анализов.

Задачи корреляционного анализа сводятся к измерению тесноты связи между факторными и результативными признаками и к оценке факторов, оказывающих наибольшее влияние на результативный признак.

Регрессионная модель. Построение регрессионной модели.

Задачи регрессионного анализа сводятся к определению типа связей между теми же объектами, что и выше, и построению функции регрессии, позволяющей вычислить среднее значение результативного признака в зависимости от значений фактора.

Показатели тесноты связи между признаками называют коэффициентами корреляции. Их выбор зависит от того, в каких шкалах измерены признаки.

1) номинальная шкала предназначена для описания принадлежности объектов к определенным социальным группам (н-р: ИТР, рабочий, мастер и т.д.)
2) шкала порядка применяется для измерения упорядоченности объектов по одному или нескольким признакам (экзаменационные оценки, тестовые баллы).
3) Количественная шкала используется для описания количественных показателей (заработная плата, численность группы и т.д.).

Выявление связи между признаками осуществляется следующим образом: выдвигается нулевая статистическая гипотеза об отсутствии связи между признаками; рассчитывается соответствующий коэффициент корреляции «к»; проверяется, превосходит ли он некоторое критическое значение $к_{крит}$. Если $к>к_{крит}$, то гипотеза об отсутствии связи отвергается.

Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции.

1. Расчет линейного коэффициента корреляции Пирсона для несгруппированных данных производят по формулам:

$$R=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum\ (y-\bar{y})^2}}$$

Где x и y – значения признаков, $\bar{x}$ u $\bar{y}$ - их средние значения;

Или

$$R=\frac{n\sum xy-\sum x\sum\ y}{\sqrt{\left[n\sum x^2-(\sum x)^2\right]\cdot\left[n\sum y^2-(\sum y)^2\right]}}$$

Где x и y – значения признаков, между которыми определяется коэффициент корреляции; n - объем выборки.

Линейный коэффициент корреляции $|R|<=1$. Знак коэффициента характеризует направление взаимосвязи. Абсолютная величина R характеризует степень тесноты рассматриваемой взаимосвязи.

Если $|R|<0.3$, то связь практически отсутствует; если $0.3<=|R|< 0.5$, то связь слабая; если $0.5<=|R|< 0.7$, то связь достаточно сильная; если $|R|>0.7$, то имеется высокая зависимость между признаками.

2. Корреляционное отношение:

$$\eta=\sqrt{\frac{\sigma_{xy}^2}{\sigma_y^2}}$$

где $\sigma_{ух}^2$ – характеризует вариацию результативного признака под влиянием вариации признака фактора, а $\sigma_у^2$ – характеризует вариацию результативного признака под влиянием всех факторов.

3. Для анализа связей между признаками, измеренных в порядковых шкалах, применяют ранговые коэффициенты корреляции.

А. Расчет рангового коэффициента корреляции Спирмена производится по формуле:

$$\rho = 1 - \frac{6}{n(n^2 - 1)} \sum_{k=1}^{n} (A_k - B_k)^2$$

Где $А_к$ – ранг к-го наблюдения по показателю х;
$В_к$ – ранг к-го наблюдения по показателю у;
n - число пар наблюдений.

Проверка статистической существенности показателя осуществляется по таблицам критических значений.

Б. Расчет рангового коэффициента корреляции Кендалла производится по формуле:

$$\tau = \frac{2S}{n(n-1)}$$

где S – сумма баллов, если баллом +1 оценивается пара рангов, имеющих по обоим показателям одинаковый порядок, а балом –1 – пара рангов с различным порядком.

4. Значительная часть признаков, измеренная с помощью номинальных шкал, составляют признаки, принимающие два и более альтернативных значения.
Для анализа таких признаков рассчитываются коэффициенты ассоциации и контингенции.
Коэффициент ассоциации:

$$k_a = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c}$$

Предельным для коэффициента ассоциации является 0.5, т.е. если $к_а$>0.5, то между признаками имеется существенная взаимосвязь.
Коэффициент контингенции:

$$k_k = \frac{ad - bc}{\sqrt{(a+d)(c+d)(a+c)(b+d)}}$$

Решение принимается следующим образом: если $к_к$>=0.3, то это свидетельствует о наличии связи между качественными признаками.

5. Если признаки имеют несколько возможных значений, то используют коэффициенты сопряженности Чупрова или Пирсона.
Расчет коэффициента Пирсона производится по формуле:

$$\Phi^2 = \sum_{i=1}^{m1} \sum_{j=1}^{m2} \frac{(\varphi_{ij} - \varphi_i \varphi_j)^2}{\varphi_i \varphi_j}$$

Где: $\varphi_j = \sum_{i=1}^{m1} \varphi_{ij}; \quad \varphi_i = \sum_{j=1}^{m2} \varphi_{ij}; \quad i = 1,..., m_1; \quad j = 1,..., m_2$

Где: φ_{ij} - доля объектов, имеющих i–е значение одного и j-е значение другого признака; m_1 – количество градаций первого признака; m_2 – количество градаций второго признака.
Коэффициент Чупрова определяется следующим образом:

$$T^2 = \frac{\Phi^2}{\sqrt{(m_1 - 1)(m_2 - 1)}}$$

Коэффициент Пирсона статистически существенен, если

$$\Phi^2 \geq \frac{\chi^2}{N}$$

Где χ^2 – табличное значения критерия кси-квадрат с $(m_1-1)(m_2-1)$ степенями свободы; N - количество наблюдений.

Явления общественной жизни складываются под воздействием не одного, а целого ряда факторов, т. е. эти явления многофакторны. Между факторами существуют сложные взаимосвязи, поэтому их влияния комплексное и его нельзя рассматривать как простую сумму изолированных влияний.

Многофакторный корреляционно-регрессионный анализ позволяет оценить меру влияния на исследуемый результативный показатель каждого из включенных в модель (уравнение) факторов при любых возможных сочетаниях факторов найти с определенной степенью точностью теоретическое значение этого показателя. Важным условием является отсутствие между факторами функциональной связи.

4.3. Лабораторные работы

№ п/п	*Номер раздела дисциплины*	*Наименование лабораторной работы*	*Объем (час.)*	*Вид занятия в интерактивной, активной, инновационной формах, (час.)*
1	**1.**	Обработка данных	4	Работа в малых группах (1час.)
2	**2.**	Статистические критерии	6	Работа в малых группах (1час.)
3	**5.**	Дисперсионный анализ	6	Работа в малых группах (1час.)
4	**6.**	Регрессионный анализ	6	Работа в малых группах (1час.)
5	**5.**	Многомерные методы. Факторный анализ.	6	Работа в малых группах (2час.)
6	**6.**	Анализ и прогноз тренда	6	Работа в малых группах (1час.)
		ИТОГО	**34**	**7**

4.4. Практические занятия

№ п/п	*Номер раздела дисциплины*	*Наименование тем практических занятий*	*Объем (час.)*	*Вид занятия в интерактивной, активной, инновационной формах, (час.)*
1	**1.**	Числовые характеристики случайной величины.	2	разбор конкретных ситуаций (2час.)
2	**2.**	Функции распределения вероятностей	3	-
3	**3**	Определение точечных оценок. Определение интервальных оценок	3	-
4	**4.**	Этапы проверки гипотез	3	-
5	**5.**	Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ.	3	-
6	**6.**	Многофакторный корреляционно-регрессионный анализ	3	-
		ИТОГО	**17**	**2**

4.5. Контрольные мероприятия: курсовой проект (курсовая работа), контрольная работа, РГР, реферат

учебным планом не предусмотрено

5. МАТРИЦА СООТНЕСЕНИЯ РАЗДЕЛОВ УЧЕБНОЙ ДИСЦИПЛИНЫ К ФОРМИРУЕМЫМ В НИХ КОМПЕТЕНЦИЯМ И ОЦЕНКЕ РЕЗУЛЬТАТОВ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Компетенции / *№, наименование разделов дисциплины*	*Кол-во часов*	*Компетенции*		*Σ комп.*	$t_{ср}$*, час*	*Вид учебных занятий*	*Оценка результатов*
		ОПК	*ПК*				
		2	**2**				
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1.Основные сведения	16	-	+	1	16	Лк,ПЗ,ЛР, СРС	Экзамен
2.Основные законы распределения вероятностей	23	-	+	1	23	Лк,ПЗ,ЛР, СРС	Экзамен
3.Методы получения точечных оценок. Интервальные оценки	20	-	+	1	20	Лк,ПЗ, СРС	Экзамен
4.Проверка статистических гипотез	18	-	+	1	18	Лк,ПЗ, СРС	Экзамен
5.Однофакторный, двухфакторный анализ	24	+	-	1	24	Лк,ПЗ,ЛР, СРС	Экзамен
6.Корреляционно- регрессионный анализ	25	+	-	1	25	Лк,ПЗ,ЛР, СРС	Экзамен
всего часов	**126**	**49**	**77**	**2**	**63**		

6. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

1 Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов / В.Е.Гмурман. - 12-е изд., перераб. – Москва : Высшее образование, 2007. - 479 с. (с.23-49, 147-210, 405-477)

7. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

№	*Наименование издания*	*Вид занятия* (*Лк, ЛР, ПЗ, кр*)	*Количество экземпляров в библиотеке,* **шт.**	*Обеспеченность,* **(экз./ чел.)**
1	**2**	**3**	**4**	**5**
	Основная литература			
1.	Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов / В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.	Лк, ПЗ, ЛР	49	1
2.	Джафаров, К.А. Теория вероятностей и математическая статистика : учебное пособие / К.А. Джафаров.-Новосибирск : Изд-во НГТУ, 2015. - 167с. //biblioclub.ru/index.php?page=book&id=438304	Лк, ПЗ	ЭР	1
	Дополнительная литература			
3.	Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.	ЛР	44	1
4.	Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с.	ПЗ	149	1

8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО - ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ» НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.Электронный каталог библиотеки БрГУ
http://irbis.brstu.ru/CGI/irbis64r_15/cgiirbis_64.exe?LNG=&C21COM=F&I21DBN=BOOK&P21DBN=BOOK&S21CNR=&Z21ID=.
2. Электронная библиотека БрГУ
http://ecat.brstu.ru/catalog .
3. Электронно-библиотечная система «Университетская библиотека online» http://biblioclub.ru .
4. Электронно-библиотечная система «Издательство «Лань»
http://e.lanbook.com .
5. Информационная система "Единое окно доступа к образовательным ресурсам"
http://window.edu.ru .
6. Научная электронная библиотека eLIBRARY.RU http://elibrary.ru .
7. Университетская информационная система РОССИЯ (УИС РОССИЯ) https://uisrussia.msu.ru/ .
8. Национальная электронная библиотека НЭБ http://xn--90ax2c.xn--p1ai/how-to-search /.

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

9.1. Методические указания для обучающихся по выполнению лабораторных работ/практических работ

<u>Лабораторная работа №1</u>

Обработка данных

<u>Цель работы:</u>

изучение способа обработки пропущенных значений в исходных данных, методов получения параметров описательной статистики, расчета распределений и частот и контроля качества.

<u>Задание:</u>

На предприятии проводились замеры изменения часовых нагрузок отопления – x_2, вентиляции – x_3 и суммарной тепловой нагрузки – x_4 в зависимости от температуры окружающего воздуха – x_1 (таблица1.1).Необходимо осуществить обработку исходных данных в зависимости от варианта согласно содержанию и порядка выполнения:

Таблица 1.1

Исходные данные для выполнения лабораторных работ № 1-4

Вариант 1				
№ п/п	**Температура воздуха, ^{0}C – x_1**	**Отопление, Гкал/ч –x_2**	**Вентиляция, Гкал/ч –x_3**	**Суммарная наг-рузка, Гкал/ч –x_4**
1.	8	0,308	0,119	0,634
2.	7	0,359	-	0,709
3.	6	0,411	0,165	-
4.	5	0,463	0,188	0,859
5.	4	-	0,211	0,934
6.	3	0,567	0,234	-
7.	2	0,618	0,257	1,083
8.	1	0,670	-	1,158
9.	0	-	0,303	1,233
10.	-1	0,774	0,326	1,308
Вариант 2				
№ п/п	**Температура воздуха, ^{0}C – x_1**	**Отопление, Гкал/ч –x_2**	**Вентиляция, Гкал/ч –x_3**	**Суммарная наг-рузка, Гкал/ч –x_4**
1.	-2	0,826	-	1,383
2.	-3	-	0,372	1,457
3.	-4	0,926	0,395	-
4.	-5	0,981	0,418	1,607
5.	-6	1,033	-	1,682
6.	-7	1,085	0,464	1,757
7.	-8	1,136	0,488	-
8.	-9	1,188	0,511	1,907
9.	-10	-	0,534	1,982
10.	-11	1,318	0,581	2,082

Вариант 3				
№ п/п	**Температура воздуха, ^{0}C – x_1**	**Отопление, Гкал/ч –x_2**	**Вентиляция, Гкал/ч –x_3**	**Суммарная наг-рузка, Гкал/ч –x_4**

1.	-12	1,344	-	2,131
2.	-13	1,395	0,603	2,206
3.	-14	1,447	0,626	-
4.	-15	1,499	-	2,359
5.	-16	1,551	0,672	2,431
6.	-17	-	0,695	2,505
7.	-18	1,654	0,718	2,580
8.	-19	1,706	0,741	2,655
9.	-20	-	0,764	2,730
10.	-21	1,81	0,787	-

Вариант 4

№ п/п	Температура воздуха, ^{0}С – x_1	Отопление, Гкал/ч –x_2	Вентиляция, Гкал/ч –x_3	Суммарная наг-рузка, Гкал/ч –x_4
1.	-22	-	0,810	3,058
2.	-23	1,913	0,833	3,110
3.	-24	1,965	-	3,162
4.	-25	-	0,857	3,213
5.	-26	2,069	0,802	-
6.	-27	2,121	0,873	3,317
7.	-28	2,172	0,885	3,369
8.	-29	2,224	0,864	-
9.	-30	2,276	0,890	3,472
10.	-31	2,328	-	3,524

Вариант 5

№ п/п	Температура воздуха, ^{0}С – x_1	Отопление, Гкал/ч –x_2	Вентиляция, Гкал/ч –x_3	Суммарная наг-рузка, Гкал/ч –x_4
1.	-5	0,981	0,418	1,607
2.	-6	1,033	-	1,682
3.	-7	-	0,464	1,757
4.	-8	1,136	0,488	-
5.	-9	1,188	0,511	1,907
6.	-10	-	0,534	1,982
7.	-11	1,318	0,581	2,082
8.	-12	1,344	-	2,131
9.	-13	1,395	0,603	2,206
10.	-14	1,447	0,626	-

Вариант 6

№ п/п	Температура воздуха, ^{0}С – x_1	Отопление, Гкал/ч –x_2	Вентиляция, Гкал/ч –x_3	Суммарная наг-рузка, Гкал/ч –x_4
1.	-15	1,499	-	2,359
2.	-16	1,551	0,672	2,431
3.	-17	-	0,695	2,505
4.	-18	1,654	0,718	2,580
5.	-19	1,706	0,741	2,655
6.	-20	-	0,764	2,730
7.	-21	1,81	0,787	-
8.	-22	-	0,810	3,058

9.	-23	1,913	0,833	3,110
10.	-24	1,965	-	3,162

Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.

Порядок выполнения:

В зависимости от варианта выбрать исходные данные. Для обработки пропущенных в них значений для четных вариантов применить метод «Средних», для нечетных – метод «Регрессии».

- По полученным исходным данным построить график «Ящики с усами».
- Для имеющихся данных рассчитать описательную статистику и на графике «Ящики с усами» отметить значения среднего и медианны.
- Построить гистограмму для всех по очереди переменных X и проверить нормальность, сделать соответствующие выводы. Если распределение не отличается от нормального, то выполняются следующие пункты задания, если же отличается, то только последний.
- Вычислить значения функции распределения вероятностей и построить график вероятности и плотности вероятности числа успехов.
- Построить для всех переменных X диаграмму Парето.

Форма отчетности:

Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе

Ознакомиться с теоретическим материалом, представленным в первом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. Что такое медиана?
2. Дать определение выборочной средней.
3. Как определяется стандартное отклонение?
4. Что такое дисперсия?
5. Что определяет критерий Колмогорова?
6. Когда применяется критерий омега-квадрат?
7. Для чего используется критерий хи-квадрат?
8. Какое распределение является нормальным?

Лабораторная работа №2
Статистические критерии

Цель работы:

изучение параметрических и непараметрических критериев статистики и их применение для анализа данных.

Задание:

В зависимости от варианта (см лабораторную работу №1) выбрать исходные данные.

Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.

Порядок выполнения:

- Используя в пункте «Статистика» процедуру «Корреляция» рассчитать парные коэффициенты корреляции, описать тесноту и направление связи между ними и выделить значимые. Сопоставить полученные результаты с функциональными графиками, построенными для каждой пары исследуемых переменных.
- Для заданных исходных данных рассчитать критерии Фишера и Стьюдента, сделать соответствующие выводы по ним.
- Преобразовать исходные выборки в форму гистограммы.
- В зависимости от анализируемых данных определить непараметрический критерий хи - квадрат для принятия заданной нулевой или альтернативной гипотезы.
- Осуществить процедуру ранговой корреляции и по полученным результатам сделать соответствующие выводы.

Форма отчетности:

Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе

Ознакомиться с теоретическим материалом, представленным во втором разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. Что относится к параметрическим процедурам?
2. Какие процедуры называются непараметрическими?
3. Что показывают парные коэффициенты корреляции?
4. Что определяют критерии Фишера и Стьюдента?
5. Что такое ранг?

Лабораторная работа №3
Дисперсионный анализ

Цель работы:

изучение методики проведения дисперсионного анализа, состоящей из однофакторного, двухфакторного и многофакторного анализа.

Задание:

В зависимости от варианта (см лабораторную работу №1) выбрать исходные данные.

Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.

Порядок выполнения:

В лабораторной работе необходимо выполнить:

- Параметрический однофакторный дисперсионный анализ.
- Параметрический двухфакторный дисперсионный анализ без повторных измерений.
- Непараметрический однофакторный дисперсионный анализ.
- Многофакторный дисперсионный анализ для исследуемых переменных.

Форма отчетности:

Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе

Ознакомиться с теоретическим материалом, представленным в пятом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. Область применения дисперсионного анализа.
2. Методика проведения дисперсионного анализа.
3. Для чего используется критерий Фишера в двухфакторном дисперсионном анализе?
4. Чем отличается эксперимент без повторных измерений от эксперимента с повторными измерениями?
5. Назначение критерия Крускала-Уоллиса.
6. Применение критерия Джонкхиера.
7. Какие отличия двухфакторного анализа от много-факторного?

Лабораторная работа №4
Регрессионный анализ

Цель работы:

изучение методов регрессионного анализа, методики построения регрессионных моделей; выполнение расчетов простой регрессии и множественной линейной регрессии.

Задание:

В зависимости от варианта (см лабораторную работу №1) выбрать исходные.

Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.

Порядок выполнения:

В лабораторной работе необходимо выполнить:

- Расчет простой регрессии для переменных X_1, X_2, X_3, X_4 по очереди, приняв переменную X_5 за Y.
- Построить для каждой переменной X_i адекватную регрессионную модель.
- Анализ регрессионных остатков.
- Сделать прогнозы для полученных адекватных моделей.
- Расчет множественной линейной регрессии.

Форма отчетности:

Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:
Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе
Ознакомиться с теоретическим материалом, представленным в шестом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2.Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. Какая модель называется линейной?
2. Что называется параболической моделью?
3. Какая зависимость называется логарифмической?
4. Что такое экспонента?
5. Виды регрессионных моделей.
6. Методика регрессионного анализа.
7. Какая модель является адекватной?
8. Какая регрессия называется простой?
9. Что называется множественной регрессией?

Лабораторная работа №5
Многомерные методы. Факторный анализ

Цель работы:

изучение методов проведения факторного анализа, методики выполнения вращения факторов.

Задание:

Заполнить исходную матрицу 5×8 (табл. 1.2), например «политические лидеры – свойства» своими собственными ответами или ответами других людей и провести факторный анализ.

Таблица 1.2

Исходные данные для лабораторной работы № 5

лидеры / свойства	x_1	x_2	x_3	x_4	x_5
1					
2					
3					
4					
5					
6					
7					
8					

Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.
Порядок выполнения:

- Расчет средних значений и корреляционную матрицу переменных.
- Нормализацию факторных нагрузок (Кайзера).
- Проекции нагрузок для главных компонент.
- Вращение факторов методом «Варимакс».

Форма отчетности:
Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:
Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе
Ознакомиться с теоретическим материалом, представленным в пятом разделе данной дисциплины.

Основная литература
1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.
Дополнительная литература
2.Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. В каких случаях осуществляется факторный анализ?
2. Методика проведения факторного анализа.
3. Что такое факторные нагрузки?
4. Основные формы вращения нагрузок.
5. В чем отличие метода вращения «варимакс» от других методов вращения?

Лабораторная работа №6
Анализ и прогноз тренда

Цель работы:
изучение методов осуществления анализа и прогноза тренда, применение корреляционного и регрессионного анализа.
Задание:
Выполнить анализ и прогноз тренда валютных котировок EUR/USD для графиков выданных преподавателем, в качестве сдвигаемого фрагмента принять 5 первых измерений
Вид занятий в интерактивной форме: выполнить задание в группе из 2-3 человек.
Порядок выполнения:
В лабораторной работе необходимо выполнить:
- В зависимости от варианта в разделе задания на лабораторную работу №6 выбрать графики валютных котировок и их средних значений.
- За указанный период на графике валютных котировок указать основные тренды.
- Составить таблицу исходных.
- Корреляционный анализ временного ряда, включающий построение автокорреляционной функции.

- Регрессионный анализ с использованием простой регрессии, включающей: построение адекватной математической модели, анализ остатков, прогнозирование;

Форма отчетности:

Отчет набирается на компьютере и сдается в печатном виде. В отчете должны присутствовать:

1. Номер варианта
2. Цель работы
3. Задание
4. Поэтапное выполнение всех заданий варианта
5. Заключение.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к лабораторной работе

Ознакомиться с теоретическим материалом, представленным в шестом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2.Сальникова М.К. Математическая статистика. Применение методов анализа данных с использованием интегрированного статистического пакета STADIA: методические указания к выполнению лабораторных работ/ М.К. Сальникова, Братск: ГОУ ВПО «БрГУ», 2006.-79с.

Контрольные вопросы для самопроверки

1. Что называется временным рядом?
2. Что такое тренд?
3. Методика проведения анализа тренда.
4. Для чего проводится корреляционный анализ временных рядов?
5. Что такое классическая корреляция?
6. Что называется интервальной корреляцией?
7. Какая функция называется автокорреляционной?

Практическое занятие № 1.
Числовые характеристики случайной величины.

Цель работы:

Изучить силовые характеристики случайных величин, решить задачи.

Вид занятий в интерактивной форме: выполнить задание и разобрать конкретный пример.

Задание:

Пример 1.

Пусть из генеральной совокупности извлечена выборка (табл.1):

Таблица 1.

Значение	2	5	7	4	2	5

Составим вариационный ряд. Он имеет вид: 2;2;4;5;5;7. Т.е. варианты записаны в возрастающем порядке.

Статистическим распределением выборки называют перечень вариант и соответствующих им частот n_i или относительных частот:

Построим статистическое распределение выборки (СРВ):

Таблица 2.

СРВ через частоты вариант

	2	4	5	7
	2	1	2	1

Таблица 3.

СРВ через относительные частоты при n_i=6 – объем выборки

	2	4	5	7
	1/3	1/6	1/3	1/6

СВР можно также задать и с помощью интервалов и соответствующих им частот. В качестве частоты интервала принимают сумму частот вариант, попавших в этот интервал.

Пример 2.

Пусть выборка задана следующим условием:

Таблица 4.

	2	3	5	6	1
	2	3	1	1	2

Представим СРВ с помощью последовательности интервалов. Для этого сначала составляется СРВ через частоты вариант, т.е. расположим варианты в возрастающем порядке и укажем их частоты:

Таблица 5.

	1	2	3	5	6
	2	2	3	1	1

-На числовой прямой расположим варианты

-Возьмем интервал необходимой длины, например: 2.1. Тогда СРВ с помощью интервалов имеет вид:

Таблица 6.

Номер интервала	Частичный интервал	Сумма частот вариант, попавших в интервал	Плотность частоты
1	(0;2.1)	4	4/2.1
2	(2.1;4.2)	3	3/2.1
3	(4.2;6.3)	2	2/2.1

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к практическим занятиям:

Ознакомиться с теоретическим материалом, представленным в первом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с.

Практическое занятие № 2

Функции распределения вероятностей.

Цель работы:

Изучить функции распределения вероятностей, научиться строить характеристики, решить задачи.

Задание:

Пример 1.

Найдем функцию распределения выборки (см. практическое занятие №1, пример 1).

Таблица 1.

Значение	2	5	7	4	2	5

$$F(x) = n_x / n,$$

где n – объем выборки, n_x – число вариант выборки, меньших x.

Число 2 является наименьшим значением вариант, а 7 – наибольшим. Тогда функция распределения будет иметь вид:

$$F(X)=\begin{cases}0, & x\le 2;\\ 1/3, & 2<x\le 4;\\ 1/2, & 4<x\le 5;\\ 5/6, & 5<x\le 7;\\ 1, & 7>x.\end{cases}$$

Построить график этой функции.

Пример 2.

Построим полигон частот и относительных частот выборки для примера 1.

Значения берутся из таблиц 2 и 3 соответственно.

Пример 3.

Построим гистограмму частот выборки для примера 2.

Таблица 8.

	2	3	5	6	1
	2	3	1	1	2

Для этого используется таблица 5.

Построить гистограмму относительных частот. Объем выборки для этого случая будет равен 9. Поэтому разделив на 9, получим искомые высоты прямоугольников.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к практическим занятиям:

Ознакомиться с теоретическим материалом, представленным во втором разделе данной дисциплины.

Основная литература

1. Джафаров, К.А. Теория вероятностей и математическая статистика : учебное пособие / К.А. Джафаров.-Новосибирск : Изд-во НГТУ, 2015. - 167с.

Дополнительная литература

2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с..

Практическое занятие № 3

Определение точечных оценок. Определение интервальных оценок.

Цель работы:

Изучить точечные и интервальные оценки, решить задачи.

Задание:

Пример 1.

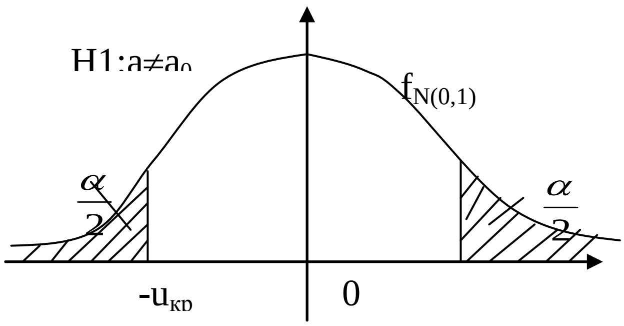
Из нормальной генеральной совокупности с известным средним квадратическим отклонением σ=5.2 извлечена выборка объема n=100 и по ней найдена выборочная средняя $\bar{x}=27.56$. Требуется а) при уровне значимости 0.05 проверить нулевую гипотезу Ho: a=a_0=26 при конкурирующей гипотезе H1: a≠26; б) найти мощность критерия.

Решение. А) найдем наблюдаемое значение критерия

$$U_{набл}=\frac{(\bar{x}-a_0)\sqrt{n}}{\sigma}=\frac{(27/56-26)\sqrt{100}}{5.2}=3.$$

По условию, конкурирующая гипотеза имеет вид a≠a_0, поэтому критическая область двусторонняя.

График двусторонней критической области в этом случае имеет вид:



Найдем критическую точку из равенства

$$\Phi(u_{кр}) = (1-\alpha)/2 = (1-0.05)/2 = 0.475$$

По таблице функции Лапласа находим $u_{кр}$=1.96.

Так как $|U_{набл}| > u_{кр}$, то нулевую гипотезу отвергаем, другими словами, выборочная и гипотетическая генеральная средние различаются значимо.

Б) Мощность критерия проверки гипотезы Но зависит от вида Н1. В нашем случае Н1: $a \neq a_0$, поэтому мощность двустороннего критерия равна

$$1-\beta = 1-((\Phi(u_{кр}-\lambda)+(\Phi(u_{кр}+\lambda)),$$

Где $u_{кр}$ и λ находят из равенств:

$$\Phi(u_{кр}) = \frac{1-\alpha}{2}, \quad \lambda = \frac{(a_1-a_0)\sqrt{n}}{\sigma}, \quad (a_1 = a = \bar{x} = 27.56; \quad a_0 = 26)$$

Тогда:

$$\Phi(u_{кр}) = \frac{1-0.05}{2} = 0.475, \quad u_{кр} = 1.96; \quad \lambda = \frac{(27.56-26)\sqrt{100}}{5.2} = 3;$$

$$1-\beta = 1-((\Phi(1.96-3)+(\Phi(1.96+3)) = 1-(\Phi(-1.04)+\Phi(4.96)) =$$

$$= 1-(-0.3508+0.5) = 0.8508$$

Итак, мощность критерия 1-β=0.8508

Пример 2.

Пусть выполнены все условия предыдущего примера, кроме альтернативной гипотезы Н1. Здесь Н1: а=27.56> а =26.

Требуется: а) проверить гипотезу Но: $a=a_0$; б) найти мощность критерия проверки Но.

Решение: а) число $U_{набл}$ находится как:

$$U_{набл} = \frac{(\bar{x}-a_0)\sqrt{n}}{\sigma} = \frac{(27.56-26)\sqrt{100}}{5.2} = 3.$$

Так как $a>a_0$, то критическая область правосторонняя, ее график имеет вид:

по правилу 2 число находится из условия:

$$\Phi(u_{кр}) = \frac{1-2\alpha}{2} = \frac{1-2*0.05}{2} = 0.45$$

Из таблицы следует, что $u_{кр}$=1.645. $U_{набл}$=3>$u_{кр}$=1.645 , поэтому гипотеза Но отвергается.

Б) При конкурирующей гипотезе Н1: $a>a_0$ мощность правостороннего критерия 1- β равна

$$1-\beta = 0.5-\Phi(u_{кр}-\lambda)$$

Где $u_{кр}$ находят также как и выше из условия:

$$\Phi(u_{кр}) = \frac{1-2\alpha}{2}, \quad \lambda = \frac{(a_1 - a_0)\sqrt{n}}{\sigma} = (27.56-26)\frac{\sqrt{100}}{5.2} = 3, \quad u_{кр} = 1.645$$

Тогда мощность критерия проверки Но равна:

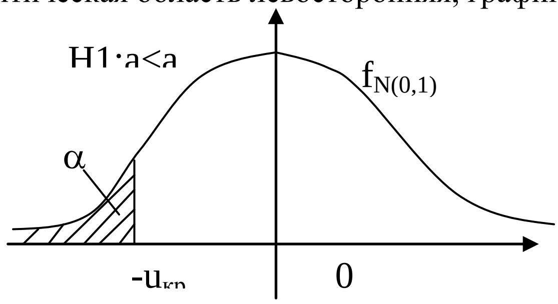
$$1-\beta = 0.5 - \Phi(1.645-3) = 0.5 + \Phi(-1.355) = 0.5 + 0.4091 = 0.9091$$

Пример 3.

По результатам n=9 замеров установлено, что среднее время изготовления детали $\bar{x}$ =48с. предполагая, что время изготовления – нормально распределенная случайная величина с дисперсией $\sigma^2=9c^2$, на уровне значимости $\alpha=0.05$ решить: а) можно ли принять 50с в качестве нормативного времени (математического ожидания) изготовления детали? б) можно ли принять за норматив 49с? Каковы вероятности того, что принятые решения будут ошибочными?

Решение. А) по условию, нулевая гипотеза Но: a_0=50с. Т.к. в результате выборочных наблюдений получено $\bar{x}$=48с, то в качестве альтернативной возьмем гипотезу Н1: а=48с<50с.

Так как $a<a_0$, то критическая область левосторонняя, график ее имеет вид:



принимая правило 3, получим: $u_{кр}^{/}=-u_{кр}$, $u_{кр}$ находим из равенства

$$\Phi(u_{кр}) = \frac{1-2\alpha}{2} = \frac{1-2*0.05}{2} = 0.45, \quad u_{кр} = 1.645$$

Следовательно,

$$u'_{кр} = -1.645$$

$$U_{набл} = \frac{(\bar{x}-a_0)\sqrt{n}}{\sigma} = \frac{(48-50)\sqrt{9}}{3} = -2.$$

Так как $U_{набл}=-2< u'_{кр}=-1.645$, то по правилу три нулевую гипотезу Но должны отвергнуть.

Б) Здесь a_0=49с, а=48с., т.е. $a_0>a$, поэтому применяем правило 3. Из предыдущего пункта следует, что $u_{кр} = 1.645$; $u'_{кр} = -1.645$

$$U_{набл} = \frac{(\bar{x}-a_0)\sqrt{n}}{\sigma} = \frac{(48-49)\sqrt{9}}{3} = -1.$$

Так как $U_{набл}>-u_{кр}$, то Но не отвергаем.

Вычислим вероятность ошибки второго рода

$$\beta = P_{H1}(Ho) = \frac{1}{2} + \frac{1}{2}\Phi(u_{кр}+\lambda),$$

где $u_{кр} = 1.645$, $\lambda = (a-a_0)\frac{\sqrt{n}}{\sigma} = \frac{(48-49)\sqrt{9}}{3} = -1$

$$p = \frac{1}{2} + \frac{1}{2}\Phi(0.645) = \frac{1}{2} + \frac{1}{2} * 0.4843 = 0.7422.$$

Итак, принимая за норматив а=49с, можно допустить ошибку второго рода, ее вероятность р=0.7422.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к практическим занятиям:

Ознакомиться с теоретическим материалом, представленным в третьем разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с..

Практическое занятие № 4

Этапы проверки гипотез.

Цель работы:

Изучить этапы проверки гипотез, решить задачи.

Задание:

Произведено по четыре испытания на каждом из трех уровней фактора F. Методом дисперсионного анализа при уровне значимости 0.05 проверить нулевую гипотезу о равенстве групповых средних. Предполагается, что выборки извлечены из нормальных совокупностей с одинаковыми дисперсиями. Результаты испытаний приведены в таблице:

Номер испытания i	Уровни фактора		
	F1	F2	F3
1	38	20	21
2	36	24	22
3	35	26	31
4	31	30	34
Групповая средняя $\bar{x}_j$			

1. Определить групповые средние.
2. Применять величину С равную групповой средней
3. Заполнить таблицу
4.

Номер испытания	УРОВНИ ФАКТОРА						Итоговый столбец
	F1		F2		F3		
	y_{i1}	y_{i1}^2	y_{i2}	y_{i2}^2	y_{i3}	y_{i3}^2	
1							-
2							-
3							-
4							-
$Q_j = \sum_{i=1}^{q} y_{ij}^2$	-		-		-		$\sum Qj =$
$T_j = \sum_{i=1}^{q} y_{ij}$		-		-		-	$\sum Tj =$
T_j^2		-		-		-	$\sum T_j^2 =$

4. Найти: $s_{общ}$, $s_{факт}$, $s_{ост}$, $s^2_{факт}$, $s^2_{ост}$, $F_{набл}$, $F_{кр}$ -?

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к практическим занятиям:

Ознакомиться с теоретическим материалом, представленным в четвертом разделе данной дисциплины.

Основная литература

1. Джафаров, К.А. Теория вероятностей и математическая статистика : учебное пособие / К.А. Джафаров.-Новосибирск : Изд-во НГТУ, 2015. - 167с.

Дополнительная литература

2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с.

Практическое занятие № 5

Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ.

Цель работы:
Изучить дисперсионный анализ, решить задачи.
Задание:
Пример 1.

Результаты зависимости частоты дрожания корпуса двигателя как функция специального браслета, одетого на корпус даны в табл.1.

Таблица 1.

Вес браслета(фунт)	0	1.25	2.5	5	7.5
Испытуемый / Обработка	1	2	3	4	5
1	3.01	2.85	2.62	2.63	2.58
2	3.47	3.43	3.15	2.83	2.70
3	3.35	3.14	3.02	2.71	2.78
4	3.10	2.86	2.58	2.49	2.36
5	3.41	3.32	3.08	2.96	2.67
6	3.07	3.06	2.85	2.50	2.43

Проверить нулевую гипотезу: Но –дрожание корпуса не зависит от веса браслета.

Пусть α - уровень значимости. Перейдем от данных табл.1 к их рангам (табл.2).

Испытуемый/обработка	1	2	3	4	5
1	5	4	2	3	1
2	5	4	3	2	1
3	5	4	3	1	2
4	5	4	3	2	1
5	5	4	3	2	1
6	5	4	3	2	1
r_j	30	24	17	12	7
$r._j$	5	4	2.8333	2	1.1667

r_j – сумма рангов по каждому столбцу; $r._j$ – средние рангов по столбцам.

Подставляя эти значения в выражение статистики Фридмана при n=6 и k=5, получим:

$$S=\left[\frac{12}{nk(k+1)}\sum_{j=1}^{k} r_j^{\,2}\right]-3n(k+1)=\frac{12}{6\cdot 5(5+1)}(30^2+24^2+16^2+12^2+7^2)-3\cdot 6(5+1)=22.5333$$

Для проверки с помощью статистики S гипотезы Но против произвольных альтернатив пользуются распределением χ^2 с (k-1) степенями свободы. При α=0.05 соответствующая процентная точка распределения есть $\chi^2(4;0.05)=9.488$. При α=0.01- χ^2 $(4;0.01)=13.292$, при α=0.001- $\chi^2(4;0.001)=18.51$. Учитывая, что $S>\chi^2(4;0.001)$, мы отвергаем гипотезу Но в пользу альтернативы на уровне значимости α=0.001. Согласно таблицам распределения $\chi^2(4)$, минимальный уровень значимости, при котором гипотеза отвергается в пользу альтернативы, равен α=0.00016.

Так как $S>\chi^2_{(1-\alpha)}(k-1)$, то гипотеза Но отвергается, принимается Н1: дрожание корпуса зависит от веса браслета.

Пример 2.

В условиях предыдущего примера за альтернативную гипотезу возьмем H1: возрастающее влияние воздействия монотонно изменяющегося уровня проверяемого фактора.

Из таблицы видно, что частота дрожания уменьшается при увеличении веса браслета. Чтобы применить критерий Пейджа необходимо обеспечить возрастающее воздействие уровней фактора. Для этого проводится перенумерация столбцов таблицы рангов в обратном порядке: номер j=1 будет соответствовать пятому столбцу, номер j=2 – четвертому столбцу и т.д.

Тогда статистика L Пейджа будет иметь вид:

$$L = \sum_{j=1}^{k} j \cdot r_j = 7 + 2*12 + 3*17 + 4*24 + 5*30 = 328$$

По таблице (прил.2) число L_{kp} при $1-\alpha = \Phi(L_{kp})$ будет равно:

При α=0.05 число L_{kp}=2.58;

Затем находим:

$$L^* = \frac{L - nk(k+1)^2/4}{\left(n\left(k^3 - k\right)^2/144(k-1)\right)^{1/2}} = \frac{328 - 6*5*(5+1)^2/4}{(6(125-5)^2/144*4)^{1/2}} = 4.75$$

Т.к. $L^* > L_{kp}$, то Ho отвергается, принимается H1.

<u>Задания для самостоятельной работы:</u>

Предусмотрены вариантом студента.

<u>Рекомендации по выполнению заданий и подготовке к практическим занятиям:</u>

Ознакомиться с теоретическим материалом, представленным в пятом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2.Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для прикладного бакалавриата / В. Е. Гмурман. - 11-е изд., перераб. и доп. – М. : Издательство Юрайт, 2017. - 404 с.

<u>Практическое занятие № 6</u>

Многофакторный корреляционно- регрессионный анализ.

<u>Цель работы:</u>

Изучить многофакторный корреляционно- регрессионный анализ, решить задачи.

<u>Задание:</u>

Пример 1.

По выборочным данным о выработке деталей за смену 20 рабочими цеха, представленными в таблице, требуется выявить зависимость производительности труда y от двух факторов: внутренних простоев x_1 и квалификации рабочих x_2.

Исходные данные для построения регрессионной модели, полученные экспериментально

Порядковый номер рабочего	Внутрисменные простои, мин x_1	Квалификация рабочего (тариф-ный разряд) x_2	Дневная выработ-ка рабочего, шт. y
1	15	3	86
2	18	4	88
3	25	5	94
4	30	2	77
5	24	4	92
Итого	112	18	437

Среднее значение	22.4	3.6	87.4

Регрессионную двухфакторную модель построим в линейной форме $\widehat{y}_{x1x2} = a_0 + a_1 x_1 + a_2 x_2$ и проверим ее адекватность.

Для нахождения параметров этого уравнения произведем вычисления вспомогательных величин и сведем их в таблицу.

y^2	x_1^2	x_2^2	yx_1	yx_2	x_1x_2
7396	225	9	1290	258	45
7744	324	16	1584	352	72
8836	625	25	2350	470	125
5929	900	4	2310	154	60
8464	576	16	2208	368	96
38369	2650	70	9742	1602	398

С учетом записи системы уравнений:

$$\begin{cases} a_0 n + a_1 \sum x_i + a_2 \sum x_2 = \sum y \\ a_0 \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1 x_2 = \sum y x_1 \\ a_0 \sum x_2 + a_1 \sum x_1 x_2 + a_2 \sum x_2^2 = \sum y x_2 \end{cases}$$

составим систему нормальных уравнений:

$$\begin{cases} 5a_0 + 112a_1 + 18a_2 = 437 \\ 112a_0 + 2650a_1 + 398a_2 = 9742 \\ 18a_0 + 398a_1 + 70a_2 = 1602 \end{cases}$$

А также определим:

$$\overline{y^2} = 7673.8; \quad \overline{x_1^2} = 530; \quad \overline{x_2^2} = 14; \quad \overline{yx_1} = 1948.4; \quad \overline{yx_2} = 320.4; \quad \overline{x_1x_2} = 79.6;$$

После построения модели необходимо проанализировать различного рода характеристики тесноты связи между зависимой и независимыми переменными и проверить адекватность модели.

1. Парные коэффициенты корреляции.

Сначала найдем средние отклонения:

$$\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{7673.8 - 7638.8} = 5.9$$

$$\sigma_{x1} = \sqrt{\overline{x_1^2} - \bar{x}_1^2} = \sqrt{530 - 501.8} = 5.3$$

$$\sigma_{x2} = \sqrt{\overline{x_2^2} - \bar{x}_2^2} = \sqrt{14 - 12.96} = 1.02$$

тогда:

$$r_{yx1} = \frac{\overline{x_1 y} - \bar{x}_1 \bar{y}}{\sigma_{x1}\sigma_y} = \frac{1948.4 - 22.4 \cdot 87.4}{5.3 \cdot 5.9} = -0.3$$

$$r_{yx2} = \frac{\overline{x_2 y} - \bar{x}_2 \bar{y}}{\sigma_{x2}\sigma_y} = \frac{320.4 - 3.6 \cdot 87.4}{1.02 \cdot 5.9} = 0.96$$

$$r_{x1x2} = \frac{\overline{x_1 x_2} - \bar{x}_1 \bar{x}_2}{\sigma_{x1}\sigma_{x2}} = \frac{79.6 - 22.4 \cdot 3.6}{5.3 \cdot 1.02} = -0.19$$

2. Частные коэффициенты корреляции:

$$r_{yx1(x2)} = \frac{r_{yx1} - r_{yx2} r_{x1x2}}{\sqrt{(1 - r_{yx2}^2)(1 - r_{x1x2}^2)}} = \frac{-0.3 - 0.96 \cdot (-0.19)}{\sqrt{(1 - (0.96)^2)(1 - (-0.19)^2)}} = -\frac{0.118}{0.275} = -0.429$$

$$r_{yx2(x1)} = \frac{r_{yx2} - r_{yx1} r_{x1x2}}{\sqrt{(1 - r_{yx1}^2)(1 - r_{x1x2}^2)}} = \frac{0.96 - (-0.3) \cdot (-0.19)}{\sqrt{(1 - (-0.3)^2)(1 - (-0.19)^2)}} = \frac{0.903}{0.9366} = 0.96$$

$$r_{x1x2(y)} = \frac{r_{x1x2} - r_{yx1} r_{yx2}}{\sqrt{(1 - r_{yx1}^2)(1 - r_{yx2}^2)}} = \frac{-0.19 - 0.96 \cdot (-0.3)}{\sqrt{(1 - (0.96)^2)(1 - (-0.3)^2)}} = \frac{0.098}{0.266} = 0.36$$

По полученным результатам можно сделать выводы: все факторы взаимосвязаны между собой и с результирующим параметром. Наиболее значащий и существенный фактор x_2 – квалификация рабочего имеет прямую зависимость с результирующим, т.е. чем выше квалификация рабочего, тем выше выработка деталей. Затем, фактор x_1 имеет обратную зависимость, т.е. чем больше простои, тем ниже выработка. При исключении фактора x_2 влияние этого фактора на результирующий усиливается.

На основе парных коэффициентов корреляции и средних отклонений можно легко рассчитать параметры уравнения линейной двухфакторной связи:

$$a_0 = \bar{y} - a_1\bar{x}_1 - a_2\bar{x}_2 = 87.4 - (-0.137)\cdot 22.4 - 5.41\cdot 3.6 = 71;$$

$$a_1 = \frac{r_{yx1} - r_{yx2}r_{x1x2}}{1 - r_{x1x2}^2}\cdot\frac{\sigma_y}{\sigma_{x1}} = \frac{-0.3 - 0.96\cdot(-0.19)}{1-(-0.19)^2}\cdot\frac{5.9}{5.3} = -0.137;$$

$$a_2 = \frac{r_{yx2} - r_{yx1}r_{x1x2}}{1 - r_{x1x2}^2}\cdot\frac{\sigma_y}{\sigma_{x2}} = \frac{0.96 - (-0.3)\cdot(-0.19)}{1-(-0.19)^2}\cdot\frac{5.9}{1.02} = 5.41$$

Уравнение множественной регрессии, выражающее зависимость производительности труда $\hat{y}$ от внутренних простоев x_1 и квалификации рабочих x_2, примет вид:

$$\hat{y}_{x1x2} = 71 - 0.137x_1 + 5.41x_2.$$

По данному уравнению заполним таблицу:

$\hat{y}_{x1x2}$	$y - \hat{y}_{x1x2}$	$(y - \hat{y}_{x1x2})^2$
85.2	0.8	0.64
90.2	-2.2	4.84
94.7	-0.7	0.49
79.1	-2.1	4.41
89.3	2.7	7.29
438.5	-	17.67

3. Совокупный коэффициент множественной корреляции:

$$R_{yx1x2} = \sqrt{\frac{r_{yx1}^2 + r_{yx2}^2 - 2r_{yx1}r_{yx2}r_{x1x2}}{1 - r_{x1x2}^2}} = \sqrt{\frac{(-0.3)^2 + 0.96^2 - 2\cdot(-0.3)\cdot 0.96\cdot(-0.19)}{1-(-0.19)^2}} = 0.967$$

4. Совокупный коэффициент множественной детерминации:

$$R^2 = 0.93$$

Данный коэффициент показывает, что вариация производительности труда на 93% обуславливается двумя анализируемыми факторами. Значит, выбранные факторы существенно влияют на показатель производительности труда. Т.о., изучаемая с помощью многофакторного корреляционного и регрессионного анализа стохастическая связь между исследуемыми показателями свидетельствует о целесообразности построения двухфакторной регрессионной модели производительности труда в виде линейного уравнения регрессии:

$$\hat{y}_{x1x2} = 71 - 0.137x_1 + 5.41x_2.$$

Проверка адекватности полученной регрессионной модели.

Проверку значимости уравнения регрессии производят на основе вычисления F – критерия Фишера:

$$F = \frac{\sigma_y^2}{\sigma_{ост}^2}\cdot\frac{n-m}{m-1} = \frac{5.9^2}{3.53}\cdot\frac{5-2}{2-1} = 29.58.$$

$$\sigma_{ост} = \sqrt{\frac{\sum(y-\hat{y})^2}{n}} = \sqrt{\frac{17.67}{5}} = \sqrt{3.53}$$

Табличное значение F – критерия при доверительной вероятности 0.95 и при числах степеней свободы k_1=m-1=2-1=1 и k_2=n-m=5-2=3 составляет 6.61. Т.к. $F_{расч}$>$F_{табл}$ больше, чем в 4 раза, данное уравнение регрессии следует признать адекватным и пригодным для практического применения.

Теперь определим значимость коэффициентов уравнения a_1 и a_2 и R_{yx1x2} с помощью t - критерия Стьюдента:

$$t_{a1} = \frac{|a_1|\sigma_{x1}\sqrt{1-r_{x1x2}^2}\cdot\sqrt{n-m-1}}{\sigma_y\sqrt{1-R_{yx1x2}^2}} = \frac{|-0.137|\cdot 5.3\cdot\sqrt{1-(-0.19)^2}\cdot\sqrt{5-2-1}}{5.9\sqrt{1-0.93^2}} = 0.459;$$

$$t_{a2} = \frac{|a_2|\sigma_{x2}\sqrt{1-r_{x1x2}^2}\cdot\sqrt{n-m-1}}{\sigma_y\sqrt{1-R_{yx1x2}^2}} = \frac{|5.41|\cdot 1.02\sqrt{1-(-0.19)^2}\cdot\sqrt{5-2-1}}{5.9\sqrt{1-0.93^2}} = 3.49$$

Существенность совокупного коэффициента корреляции определяют по формуле:

$$t_{Ryx1x2} = \frac{|R_{yx1x2}|\sqrt{n-m-1}}{1-R^2_{yx1x2}} = \frac{|0.966|\sqrt{2}}{1-0.93} = 19.32.$$

Табличное значение t – критерия при доверительной вероятности 0.9 при n-m-1=5-2-1=2 степенях свободы составляет 2.92. Величины a_2 и R_{yx1x2} следует признать значимыми, существенными.

Т.о. построенную модель пригодна для практического применения и может быть использована для выявления резервов повышения производительности труда.

Задания для самостоятельной работы:

Предусмотрены вариантом студента.

Рекомендации по выполнению заданий и подготовке к практическим занятиям:

Ознакомиться с теоретическим материалом, представленным в шестом разделе данной дисциплины.

Основная литература

1. Гмурман, В.Е. Теория вероятностей и математическая статистика : учеб. пособие для вузов/ В.Е.Гмурман. - 12-е изд., перераб. – Москва: Высшее образование, 2007. - 479 с.

Дополнительная литература

2. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - Москва : Юрайт, 2011. - 404 с.

10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

-Microsoft Imagine Premium
-ОС Windows 7 Professional
-Microsoft Office 2007 Russian Academic OPEN No Level
-Антивирусное программное обеспечение KasperskySecurity

При реализации дисциплины применяются инновационные технологии обучения, активные и интерактивные формы проведения занятий, указанные в разделах 4.3, 4.4.

11. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Вид занятия	***Наименование аудитории***	***Перечень основного оборудования***	***№ ЛР,Лк,ПЗ***
1	**3**	**4**	**5**
Лк	Дисплейный класс	AMD Athlon 64 (5GHz/250Gb/2Gb/DD-RW), 2 ядра	Лк 1-8
ЛР	Дисплейный класс	AMD Athlon 64 (5GHz/250Gb/2Gb/DD-RW), 2 ядра	ЛР 1-6
ПЗ	Дисплейный класс	AMD Athlon 64 (5GHz/250Gb/2Gb/DD-RW), 2 ядра	ПЗ 1-6
СР	Читальный зал № 3	Оборудование 15- CPU 5000/RAM 2Gb/HDD (Монитор TFT 19 LG 1953S-SF);принтер HP LaserJet P3005	

Приложение 1

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

1. Описание фонда оценочных средств (паспорт)

№ компетенции	Элемент компетенции	Раздел	Тема	ФОС
ПК-2	способность проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления	**1.**Основные сведения	1.1.Введение. Случайные величины. 1.2.Числовые характеристики случайной величины 1.3.Статистическое распределение выборки	Экзаменацион-ный билет
		2.Основные законы распределения вероятностей	2.1.Биноминальное распределение. Распределение Пуассона. Показательное распределение. 2.2.Нормальное распределение. Распределение кси-квадрат. 2.3.Распределение Стьюдента. F-распределение. Статистические оценки	Экзаменацион-ный билет
		3.Методы получения точечных оценок. Интервальные оценки	3.1.Метод максимального правдоподобия. Метод наименьших квадратов. 3.2.Интервальные оценки.	Экзаменацион-ный билет
		4.Проверка статистических гипотез	4.1. Этапы проверки гипотез. Проверка гипотезы о равенстве выборочной средней. 4.2. Проверка гипотезы о значении мат. ожидании. Проверка гипотезы о значении дисперсии.	Экзаменацион-ный билет

ОПК-2	способность выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, привлекать для их решения соответствующий физико-математический аппарат	**5**.Однофакторный, двухфакторный анализ	5.1.Виды зависимостей между признаками. Однофакторный дисперсионный анализ. 5.2. Двухфакторный дисперсионный анализ	Экзаменационный билет
		6.Корреляционно-регрессионный анализ	6.1.Многофакторный корреляционно-регрессионный анализ. Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции 6.2.Регрессионная модель. Построение регрессионной модели.	Экзаменационный билет

2. Экзаменационные вопросы

№ п/п	Компетенции		ЭКЗАМЕНАЦИОННЫЕ ВОПРОСЫ	№ и наименование раздела
	Код	Определение		
1	2	3	4	5
1.	ПК-2	способность проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления	1. Случайные величины. 2.Числовые характеристики случайной величины 3.Статистическое распределение выборки	**1.**Основные сведения
			4.Биноминальное распределение. 5.Распределение Пуассона. 6.Показательное распределение. 7.Нормальное распределение. 8.Распределение кси-квадрат. 9..Распределение Стьюдента. 10. F-распределение. 11.Статистические оценки	**2**.Основные законы распределения вероятностей
			12.Метод максимального правдоподобия. 13.Метод наименьших квадратов. 14..Интервальные оценки.	**3**.Методы получения точечных оценок. Интервальные оценки

			15.Этапы проверки гипотез. 16. Проверка гипотезы о равенстве выборочной средней. 17. Проверка гипотезы о значении мат. ожидании. 18.Проверка гипотезы о значении дисперсии.	**4**.Проверка статистических гипотез
2.	ОПК-2	способность выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, привлекать для их решения соответствующий физико-математический аппарат	19.Виды зависимостей между признаками. 20.Однофакторный дисперсионный анализ. 22. Двухфакторный дисперсионный анализ	**5**.Однофакторный, двухфакторный анализ
			23.Многофакторный корреляционно-регрессионный анализ. 24.Парные коэффициенты корреляции. 25.Частные коэффициенты корреляции. 26.Совокупные коэффициенты корреляции 27.Регрессионная модель. 28.Построение регрессионной модели	**6**.Корреляционно-регрессионный анализ

3. Описание показателей и критериев оценивания компетенций

Показатели	Оценка	Критерии
Знать **ОПК-2**: -Способность выявлять естественнонаучную сущность процесса, - Физико-математические закономерности процессов, -Способы оценки данных. **ПК-2**: - Основные приемы обработки данных, -Основные законы распределения вероятностей, -Методы обработки данных, -Способы представления экспериментальных данных. **Уметь:** **ОПК-2**: -Получать необходимую	**отлично**	Оценка «отлично» выставляется в случае, если студент демонстрирует: – всестороннее систематическое знание программного материала; – правильное выполнение практических заданий, направленных на применение программного материала; – правильное применение основных положений программного материала.
	хорошо	Оценка «хорошо» выставляется в случае, если студент демонстрирует: – недостаточно полное знание программного материала; – выполнение с несущественными ошибками практических заданий, направленных на применение программного материала; применение с несущественными ошибками основных положений программного материала.

информацию, обрабатывать ее, - Использовать полученные знания на практике, -Проводить регрессионный, дисперсионный, многофакторный корреляционно-регрессионный анализ полученных данных. **ПК-2***:* - Проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления **Владеть:** **ОПК-2***:* - Достаточным уровнем понимания материала, и способностью выявлять сущность проблем, - Соответствующим физико-математическим аппаратом для решения практических задач. **ПК-2***:* -Способностью применять стандартные программные средства с целью получения математических моделей	**удовлетворительно**	Оценка «удовлетворительно» выставляется в случае, если студент демонстрирует: – частичное знание программного материала; – частичное выполнение практических заданий, направленных на применение программного материала; частичное применение основных положений программного материала.
	неудовлетворительно	Оценка «неудовлетворительно» выставляется в случае, если студент демонстрирует: – существенные пробелы в знании программного материала; – принципиальные ошибки при выполнении практических заданий, направленных на применение программного материала; невозможность применения основных положений программного материала.

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и опыта деятельности

Дисциплина Математическая статистика направлена на ознакомление с наукой математикой, и её практическим применением в измерительных системах; на получение теоретических знаний и практических навыков использования анализами обработки данных для их дальнейшего использования в практической деятельности.

Изучение дисциплины информатика предусматривает:

– лекции,
– лабораторные работы,
– практические занятия.
– самостоятельную работу,
– экзамен.

В ходе освоения раздела 1 «Основные сведения» обучающиеся должны уяснить: разновидности случайных величин, их числовые характеристики, понятие статистического распределения выборки,

В ходе освоения раздела 2 «Основные законы распределения вероятностей» обучающиеся должны знать законы распределения вероятностей: Биноминальное распределение,

распределение Пуассона, Показательное распределение, Нормальное распределение, распределение кси-квадрат, распределение Стьюдента, F-распределение их свойства, характеристики,.

В ходе освоения раздела 3 «Методы получения точечных оценок. Интервальные оценки» обучающиеся должны знать: определения точечных, интервальных оценок, их нахождение.

В ходе освоения раздела 4 «Проверка статистических гипотез» обучающиеся должны уяснить: этапы проверки гипотез, примеры проверки гипотез.

В ходе освоения раздела 5 «Однофакторный, двухфакторный анализ» обучающиеся должны уяснить: дисперсионный анализ, уметь проводить однофакторный, двухфакторный анализ .

В ходе освоения раздела 6 «Корреляционно- регрессионный анализ» обучающиеся должны уяснить: регрессионный анализ, уметь проводить многофакторный корреляционно-регрессионный анализ.

В процессе проведения лабораторных работ происходит закрепление знаний, формирование умений и навыков реализации представления о различных способах анализа и обработки данных, построения регрессионных моделей.

В процессе проведения практических занятий вырабатывается умение использовать полученные знания на практике.

Работа с литературой является важнейшим элементом в получении знаний по дисциплине. Прежде всего, необходимо воспользоваться списком рекомендуемой по данной дисциплине литературой. Дополнительные сведения по изучаемым темам можно найти в периодической печати и Интернете.

К экзамену допускаются студенты, которые выполнили и оформили все лабораторные работы, практические занятия.

По итогам выполненного задания преподаватель оценивает уровень знаний, умений, навыков. Описание показателей и критериев оценивания компетенций, сформированных по итогам изучения дисциплины, представлено в разделе 3 Приложения 1 настоящей рабочей программы. Основными оценочными средствами при проведении промежуточной аттестации являются вопросы к экзамену.

Приложение 2

АННОТАЦИЯ
рабочей программы дисциплины
Математическая статистика

1. **Цель и задачи дисциплины**

Целью изучения дисциплины является: формирование у обучающихся знаний о сборе, обработке и представлении экспериментальных данных.

Задачей изучения дисциплины является: формирование у обучающихся знаний, умений, навыков использования основных приемов сбора, обработки и анализа данных, полученных в ходе профессиональной деятельности.

2. **Структура дисциплины**

2.1 Распределение трудоемкости по отдельным видам учебных занятий, включая самостоятельную работу: лекций – 17 часов, лабораторные работы – 34 часа, практические занятия- 17 часов, самостоятельная работа студента – 58 часов,

Общая трудоемкость дисциплины составляет 180 часов, 5 зачетных единиц.

2.2 Основные разделы дисциплины:

1. Основные сведения Количество и качество информации
2. Основные законы распределения вероятностей
3. Методы получения точечных оценок. Интервальные оценки
4. Проверка статистических гипотез
5. Однофакторный, двухфакторный анализ
6. Корреляционно- регрессионный анализ.

3. Планируемые результаты обучения (перечень компетенций)

Процесс изучения дисциплины направлен на формирование компетенции:
ОПК-2 - способность выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности, привлекать для их решения соответствующий физико-математический.
ПК-2 - способность проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления

4. Вид промежуточной аттестации: экзамен

Приложение 3

Протокол о дополнениях и изменениях в рабочей программе
на 20___-20___ учебный год

1. В рабочую программу по дисциплине вносятся следующие дополнения:
____
____

2. В рабочую программу по дисциплине вносятся следующие изменения:

____
____
____

Протокол заседания кафедры №____ от «___» __________ 20 ____ г.,
(разработчик)

Заведующий кафедрой _____________________ _____________________
(подпись) *(Ф.И.О.)*

Приложение 4

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ ПО ДИСЦИПЛИНЕ

1. Описание фонда оценочных средств (паспорт)

№ компетенции	Элемент компетенции	Раздел	Тема	ФОС
1.	ПК-2	**1.**Основные сведения	Числовые характеристики случайной величины Статистическое распределение выборки	Отчеты по лабораторным работам
		2.Основные законы распределения вероятностей	Нормальное распределение. Статистические оценки	Отчеты по лабораторным работам
2.	ОПК-2	**5**.Однофакторный, двухфакторный анализ	Виды зависимостей между признаками. Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ	Отчеты по лабораторным работам
		6.Корреляционно-регрессионный анализ	Многофакторный корреляционно-регрессионный анализ. Парные коэффициенты корреляции. Частные коэффициенты корреляции. Совокупные коэффициенты корреляции Регрессионная модель. .Построение регрессионной модели	Отчеты по лабораторным работам

2. Описание показателей и критериев оценивания компетенций

Показатели	Оценка	Критерии
Знать **ОПК-2***:* -Способность выявлять естественнонаучную сущность процесса, - Физико-математические закономерности процессов, -Способы оценки данных. **ПК-2***:* - Основные приемы обработки данных, -Основные законы распределения вероятностей, -Методы обработки данных, -Способы представления экспериментальных данных. **Уметь:** **ОПК-2***:* -Получать необходимую информацию, обрабатывать ее, - Использовать полученные знания на практике, -Проводить регрессионный, дисперсионный, многофакторный корреляционно-регрессионный анализ полученных данных. **ПК-2***:* - Проводить вычислительные эксперименты с использованием стандартных программных средств с целью получения математических моделей процессов и объектов автоматизации и управления **Владеть:** **ОПК-2***:* - Достаточным уровнем понимания материала, и способностью выявлять сущность проблем, - Соответствующим физико-математическим аппаратом для решения практических задач. **ПК-2***:* -Способностью применять стандартные программные средства с целью получения математических моделей.	**зачтено**	Оценка «зачтено» выставляется в случае, если студент демонстрирует: – всестороннее систематическое знание программного материала; – правильное выполнение лабораторных работ, направленных на применение программного материала; - правильное применение основных положений программного материала.
	не зачтено	Оценка «незачтено» выставляется в случае, если студент демонстрирует: – существенные пробелы в знании программного материала; – принципиальные ошибки при выполнении лабораторных работ, направленных на применение программного материала; - невозможность применения основных положений программного материала.