

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Самарский государственный медицинский университет»
Министерства здравоохранения Российской Федерации

Кафедра управления и экономики фармации

СОГЛАСОВАНО
Проректор по учебно-
методической работе и связям с
общественностью
профессор Т.А. Федорина

«16» 08 2016 г.

УТВЕРЖДАЮ
Председатель ЦКМС
первый проректор - проректор
по учебно-воспитательной
и социальной работе
профессор Ю.В. Щукин



«26» 08 2016 г.

РАБОЧАЯ ПРОГРАММА

обязательной дисциплины МЕДИЦИНСКАЯ СТАТИСТИКА

Шифр дисциплины: П.1.В.4

Рекомендуется для направления подготовки
31.06.00 – Клиническая медицина

Уровень образования:

высшее образование – подготовка научно-педагогических кадров в аспирантуре

Квалификация (степень) выпускника:

Исследователь. Преподаватель-исследователь

Форма обучения: **заочная**

СОГЛАСОВАНО
Проректор по научной и
инновационной работе
д.м.н. профессор
И.Л. Давыдкин

«26» 08 2016 г.

СОГЛАСОВАНО
Председатель методической
комиссии
д.м.н. Ю.В. Тезиков

«30» 08 2016 г.

Программа рассмотрена и
одобрена на заседании
кафедры управления и
экономики фармации
(протокол
№ 10 от «27» 08 2016 г.)
Заведующий кафедрой,
доцент И.К. Петрухина

«27» 08 2016 г.

Самара 2016

Программа Итоговой государственной аттестации разработана в соответствии с ФГОС ВО по направлению подготовки **31.06.00 – Клиническая медицина**, утверждённому приказом Минобрнауки РФ № 1200 от 03.09.2014, с изменениями и дополнениями от 30.04.2015; паспортом научной специальности разработанным экспертным советом Высшей аттестационной комиссии Министерства в связи с утверждением приказом Минобрнауки России от 25 февраля 2009 г. N 59 Номенклатуры специальностей научных работников (редакция от 11 ноября 2011 г.), профессиональным стандартом «Педагог профессионального обучения, профессионального образования и дополнительного профессионального образования», утвержденного приказом Министерства труда и социальной защиты от 8.09.2015 г. № 608-н; рядом других нормативных документов.

Составители рабочей программы:

Гладунова Елена Павловна, доктор фармацевтических наук, доцент, доцент кафедры управления и экономики фармации

Рецензенты:

Заведующий кафедрой управления и экономики фармации и фармацевтической технологии Государственного бюджетного образовательного учреждения высшего профессионального образования «Нижегородская медицинская академия» Министерства здравоохранения Российской Федерации, доктор фармацевтических наук **Кононова С.В.**

Доцент кафедры управления и экономики фармации, фармацевтической технологии и фармакогнозии Государственного бюджетного образовательного учреждения высшего профессионального образования «Оренбургский государственный медицинский университет» Министерства здравоохранения Российской Федерации, кандидат фармацевтических наук, доцент **Дударенкова М.Р.**

1. Цели и задачи дисциплины:

Цель - формирование у аспирантов достаточных знаний по вопросам информатики и медицинской статистики, достижение умения эффективного использования полученных знаний при выполнении научно-исследовательской работы, а так же при решении практических задач врачебной практики.

Задачи:

- формирование у аспирантов знаний о современных средствах математической статистики для решения практических задач;
- формирование понятия о специальных информационных технологиях и системах;
- формирование у аспирантов современных знаний о применении статистического метода в научных исследованиях;
- выработка умения составления программы статистического исследования;
- формирование овладения современными технологиями разработки и анализа данных на персональном компьютере, необходимыми для продуктивной научно-исследовательской работы;
- формирование умения представления статистических данных для последующей публикации результатов научных исследований

2. Место дисциплины в структуре образовательной программы

Дисциплина «Математическая статистика» относится к профессиональному циклу дисциплин (блок П.1, вариативная часть).

Дисциплина «Медицинская статистика» входит в раздел Блок 1 «Дисциплины» образовательной программы, относится к обязательным дисциплинам вариативной части, (шифр дисциплины П.1.В.4.).

Дисциплина преподается в 2-м семестре 2-го года обучения у аспирантов очной и заочной форм обучения.

Данная дисциплина базируется на знаниях, умениях и компетенциях, полученных обучающимся в процессе обучения в высшем учебном заведении, в соответствии с федеральными государственными образовательными стандартами высшего образования по программам специалиста «Лечебное дело», «Педиатрия», «Стоматология».

Изучению дисциплины в аспирантуре предшествуют «Иностранный язык», «История и философия науки», «Основы подготовки медицинской диссертации», «Психология и педагогика высшей школы».

Знания и навыки, полученные аспирантами при изучении данной дисциплины, необходимы для написания научно-квалификационной работы (диссертации); а также при подготовке к преподавательской деятельности по дисциплине «Организация фармацевтического дела».

3. Планируемые результаты обучения по дисциплине «Математическая статистика»

Процесс изучения дисциплины направлен на формирование следующих компетенций:

профессиональные компетенции:

профессиональные компетенции:

- способностью и готовностью к анализу, обобщению и публичному представлению результатов выполненных научных исследований (ОПК-3):

аспирант должен знать: информатику, статистику;

аспирант должен уметь: использовать в познавательной и профессиональной деятельности знания в области информатики, статистики;

аспирант должен владеть: элементами естественнонаучного и математического знания.

В результате изучения модуля аспирант должен:

Знать:

- современные компьютерные технологии в приложении к решению задач здравоохранения и фармации;
- методологические подходы к формализации и структуризации различных типов медицинских и фармацевтических данных для формирования решений в здравоохранении и фармации;
- виды, структуру, характеристики медицинских и фармацевтических информационных систем; принципы автоматизации управления в медико-биологических системах с использованием современных компьютерных технологий.
- определения и понятия математической статистики; значение статистического метода при проведении медико-социальных, медико-биологических исследований;
- этапы организации статистического исследования и их содержание;
- описательные статистики (виды статистических величин, методы их расчета, характеристики распределения признака в статистической совокупности, репрезентативности, среднего уровня и вариабельности данных).
- основные параметрические и непараметрические методы оценки достоверности различий статистических величин;
- основные параметрические и непараметрические методы оценки взаимосвязи между признаками;
- методы оценки динамики явлений и прогнозирования;
- метод графического изображения статистических данных;
- правила представления статистических данных для научной публикации.

Аспирант должен уметь:

разрабатывать структуры и формировать базы данных и знаний для медико-биологических систем;

проводить текстовую и графическую обработку документов с использованием стандартных программных средств;
 готовить план и программу статистического исследования;
 формировать электронную базу данных для хранения и последующей разработки данных;
 рассчитывать описательные статистики;
 строить таблицы частот и таблицы сопряженности;
 проводить оценку достоверности различий статистических величин при помощи параметрических и непараметрических методов;
 проводить дисперсионный, корреляционный и регрессионный анализы
 проводить факторный анализ;
 анализировать динамические ряды и осуществлять прогнозирование дальнейших тенденций;
 представлять статистические данные в виде графического изображения;
 представлять статистические данные для научной публикации.

Аспирант должен владеть навыками:

- основными знаниями, навыками и приемами математической статистики; статистическими методами для проведения медико-социальных, медико-биологических исследований;
- ключевыми приемами и методиками теории вероятности и математической статистики для применения их в дальнейшей научной деятельности;
- методами разработки программы научных исследований;
- методами количественной оценки случайных явлений, содержательной интерпретации полученных результатов;
- навыками оценки достоверности различий статистических величин при помощи параметрических и непараметрических методов;
- навыками проведения дисперсионного, корреляционного и регрессионного, факторного анализов.

4. Объем дисциплины и виды учебной работы

Общая трудоемкость составляет 3 зачетные единицы.

Структурные элементы программы модуля	Всего часов
Аудиторные занятия (всего)	24
Лекции	8
Практические занятия (ПЗ)	16
Семинары (С)	-
Лабораторные работы (ЛР)	-
Самостоятельная работа (всего)	82
<i>Курсовая работа</i>	-
<i>Реферат</i>	-
<i>Другие виды самостоятельной работы: домашняя подготовка к практическим</i>	

<i>занятиям и лекциям</i>	
Вид промежуточной аттестации (зачет, экзамен)	
Общая трудоемкость:	108 часов
часов	108 ч.
зачетных единиц	3 ед.

5. Содержание дисциплины, структурированное по разделам (темам) с указанием количества часов и видов занятий:

5.1. Содержание разделов дисциплины

№ раздела	Наименование раздела дисциплины	Содержание раздела	Коды компетенций
1	2	3	4
1	<i>Медицинская статистика</i>	<p>Понятие статистического эксперимента. Элементарные исходы (элементарные события). Сложные события. Частота события. События невозможные, случайные, достоверные.</p> <p>Операции над событиями.</p> <p>Объединение, пересечение дополнение. Свойства операций над событиями. Принцип двойственности.</p> <p>Свойства частот. Частота объединения и пересечения событий. Понятие условной частоты события. Независимые события (интуитивное определение). Явление статистической устойчивости частот. Введение понятия вероятности события как идеализированной "неслучайной" частоты события.</p> <p>Основные понятия математической статистики: генеральная совокупность, выборка, случайный выбор. Задача индуктивного статистического вывода – формулирование суждений о генеральной совокупности на основе выборки, извлеченной из нее случайным образом.</p>	ОПК-3
		<p>Определение вероятности события для конечного числа равновероятных (симметричных) элементарных исходов. Условная вероятность. Примеры подсчета общего числа элементарных исходов и "благоприятного" числа элементарных исходов.</p> <p>Простейшие понятия комбинаторики. Принцип сложения и принцип умножения. Сочетания и размещения. Перестановки. Выбор объектов с возвращением и без. Подсчет числа сочетаний и размещений для выбора с возвращением и без возвращения.</p> <p>Структура вероятностного пространства –</p>	ОПК-3

		<p>элементарные исходы, алгебра событий, вероятность – как функция, заданная для каждого события. Свойства вероятности. Примеры: конечное число не равновероятных элементарных исходов, бесконечное число элементарных исходов при геометрическом определении вероятности.</p> <p>Вероятность объединения событий в общем случае. Частные случаи: несовместные события, независимые события.</p> <p>Вероятность произведения событий.</p> <p>Частные случаи – независимые события, события образующие Марковскую цепь.</p> <p>Формула полной вероятности.</p> <p>Формула Байеса.</p>	
		<p>Три основных вида случайных величин – дискретные, непрерывные, смешанные. Индикатор события. Аналогия с распределением единичной массы по вещественной прямой. Атом вероятности.</p> <p>Способы задания одномерной случайной величины: ряд распределения (для дискретной с.в.), функция распределения (для любой с.в.), плотность вероятности (для непрерывной с.в.).</p> <p>Связь плотности вероятности и функции распределения ("накопленной вероятности"). Их свойства. Эмпирические аналоги функции распределения ("накопленная частота") и плотности вероятности (гистограмма).</p> <p>Среднее значение случайной величины и функции от нее – математическое ожидание.</p> <p>Моменты одномерной случайной величины – начальные и центральные. Связи между ними.</p> <p>Дисперсия (вариация). Безразмерные величины – коэффициенты вариации, асимметрии, эксцесса.</p> <p>Квантили. Медиана, квартили. Межквартильный разброс.</p> <p>Характеристики положения и рассеяния. Преимущества и недостатки использования пар - математического ожидания и среднего-квадратичного отклонения по сравнению с медианой и межквартильным разбросом.</p> <p>Производящая и характеристические функции.</p>	ОПК-3
		<p>Основные законы распределения.</p> <p>Биномиальный закон распределения. Закон распределения Пуассона. Геометрическое распределение. Равномерный закон распределения. Показательный (экспоненциальный) закон распределения.</p> <p>Нормальный закон распределения.</p> <p>Распределение некоторых случайных величин,</p>	ОПК-3

		<p>представляющих функции нормальных величин. Закон больших чисел. Неравенство Маркова (лемма Чебышева). Неравенство Чебышева. Теорема Чебышева. Теорема Бернулли. Центральная предельная теорема.</p>	
		<p>Вариационные ряды и их характеристики. Средние величины. Показатели вариации. Упрощенный способ расчета средней арифметической и дисперсии. Начальные и центральные моменты вариационного ряда.</p>	ОПК-3
		<p>Основы выборочного метода. Общие сведения о выборочном методе. Понятие оценки параметров. Оценка параметров генеральной совокупности по собственно-случайной выборке. Понятие интервального оценивания. Доверительная вероятность и предельная ошибка выборки. Оценка характеристик генеральной совокупности по малой выборке. Оценка вероятности по частоте появления события, или оценка доли объектов в генеральной совокупности по их доле в выборке, или оценка параметра биномиального распределения. Интервал рассеяния и доверительный интервал. Приближенные и точные формулы для границ доверительного интервала. Планирование объема выборки для оценки вероятности при заданных значениях точности и надежности. Понятие о принципе максимального правдоподобия на примере оценки параметра биномиального распределения. Функция распределения и плотность вероятности системы двух и более случайных величин (случайного вектора). Числовые характеристики случайных векторов: вектор математических ожиданий и матрица ковариаций. Теоремы о математическом ожидании и дисперсии. Полиномиальное распределение. Нормальное распределение для случайного вектора (на примере двумерного нормального распределения). Эллипсы рассеяния, расстояние Махаланобиса, условные плотность вероятности, математическое ожидание и дисперсия. Неравенство Чебышева. Закон больших чисел. Центральная предельная теорема Ляпунова (для частного случая: одинаково распределенных слагаемых). Распределение хи-квадрат для разных чисел степеней свободы.</p>	ОПК-3

	<p>Распределение Стьюдента. Распределение Фишера.</p>	
	<p>Основные методы построения точечных оценок – метод моментов, метод максимального правдоподобия.</p> <p>Примеры построения оценок параметров для биномиального, пуассоновского, экспоненциального распределений. Интервалы рассеяния и доверительные интервалы.</p> <p>Понятие опорной случайной величины и метод "стьюдентизации".</p> <p>Точные методы оценок параметров для нормального распределения ("теория малых выборок Стьюдента").</p> <p>Примеры проверки гипотез о параметрах распределений. Сравнение средних и дисперсий для параметров нормального распределения.</p> <p>Простые и сложные гипотезы.</p> <p>Расстояние Пирсона и критерий хи-квадрат для проверки простых и сложных гипотез.</p> <p>Критерий Колмогорова для проверки простой гипотезы о виде распределения одномерной непрерывной случайной величины.</p> <p>Выбор между двумя альтернативными гипотезами. Ошибки первого и второго рода. Мощность критерия. Случай простых гипотез – лемма Неймана-Пирсона. Критерий отношения правдоподобия.</p>	ОПК-3
	<p>Дисперсионный анализ и её сущность. Общие предпосылки использования дисперсионного анализа.</p> <p>Градации факторов и их характер. Схема варьирования при различии по одному фактору. Разное варьирование вариантов и его характеристика. Суммы квадратов и их вычисление. Степени свободы. Общая схема дисперсионного анализа при различии по одному фактору. Схема варьирования при различии по двум факторам. Суммы квадратов степени свободы и их вычисление при двух факторах. Общая схема дисперсионного анализа при различии по двум факторам.</p> <p>Пакеты статистических программ и работа с ними.</p>	ОПК-3
	<p>Виды связей и методы их анализа в статистике. Понятие о корреляционной связи. Возможности и задачи приемов сопоставления параллельных рядов, метода аналитических группировок, разложения составных показателей, индексного, графического, табличного корреляционно-регрессионного</p>	ОПК-3

		методов. Основные этапы корреляционно-регрессионного анализа. Логическая связь между признаками. Ложная корреляция. Требования к совокупности и факторным признакам, включаемым в уравнение. Определение формы связи между признаками и наличия эффекта взаимодействия факторов. Определение параметров уравнения связи, их интерпретация. Парная и множественная линейная корреляция. Коэффициенты полной и чистой регрессии. Стандартизированные коэффициенты регрессии. Криволинейная регрессия. Показатели тесноты связи. Ранговая корреляция.	
--	--	---	--

5.2. Разделы дисциплины и трудоемкость по видам учебных занятий

№ раздела	Наименование раздела дисциплины	Виды учебной работы			Всего часов
		аудиторная		внеаудиторная	
		лекции	практические занятия	СРС	
1	Медицинская статистика	36	36	36	108
	ВСЕГО:				108

6. Тематический план лекций

№ п/п	Содержание	Трудоемкость (час.)
Лекции		
1	Л.1. Случайные величины и их характеристики. Свойства математического ожидания и дисперсии.	1
2	Л.2. Вариационные ряды и их характеристики	1
3	Л.3. Основы выборочного метода	1
4	Л.4. Проверка статистических гипотез	2
5	Л.6. Дисперсионный анализ	1
6	Л.6. Корреляционный -регрессионный анализ	2
ВСЕГО:		8 часов

6. Тематический план практических занятий

№ п/п	Содержание	Формы текущего контроля	Трудоемкость (час.)
Практические занятия			
1	ПЗ 1. Теоремы сложения и умножения вероятностей. Теорема сложения вероятностей. Условная вероятность события. Теорема умножения вероятностей. Независимые события. Формула полной вероятности, формула Байеса. Повторные независимые испытания. Формула Бернулли.	<i>Устный опрос. Тестовый контроль</i>	2

	Формула Пуассона. Локальная и интегральная формулы Муавра–Лапласа.		
2	ПЗ 2. Случайные величины и их характеристики. Понятие случайной величины. Закон распределения дискретной случайной величины Математические операции над случайными величинами. Математическое ожидание дискретной случайной величины. Дисперсия дискретной случайной величины. Функция распределения случайной величины. Непрерывные случайные величины. Плотность вероятности. Мода и медиана. Квантили. Моменты случайных величин. Асимметрия и эксцесс. Свойства математического ожидания и дисперсии.	<i>Устный опрос. Тестовый контроль</i>	2
3	ПЗ 3. Вариационные ряды и их характеристики. Средние величины. Показатели вариации. Упрощенный способ расчета средней арифметической и дисперсии. Начальные и центральные моменты вариационного ряда.	<i>Устный опрос. Тестовый контроль</i>	2
4	ПЗ 4. Основы выборочного метода. Общие сведения о выборочном методе. Понятие оценки параметров. Оценка параметров генеральной совокупности по собственно-случайной выборке. Понятие интервального оценивания. Доверительная вероятность и предельная ошибка выборки. Оценка характеристик генеральной совокупности по малой выборке.	<i>Устный опрос. Тестовый контроль</i>	2
5	ПЗ 5. Дисперсионный анализ. Однофакторный дисперсионный анализ. Понятие о двухфакторном дисперсионном анализе.	<i>Устный опрос. Тестовый контроль</i>	2
6	ПЗ 6. Корреляционный анализ. Функциональная, статистическая и корреляционная зависимости. Линейная парная регрессия. Коэффициент корреляции. Основные положения корреляционного анализа. Двумерная модель. Проверка значимости и интервальная оценка параметров связи. Корреляционное отношение и индекс корреляции. Понятие о многомерном корреляционном анализе.	<i>Устный опрос. Тестовый контроль</i>	3
7	ПЗ 7. Регрессионный анализ. Основные положения регрессионного анализа. Парная регрессионная модель. Интервальная оценка и проверка значимости уравнения регрессии. Нелинейная регрессия. Множественный регрессионный анализ. Проверка значимости уравнения множественной регрессии.	<i>Устный опрос. Тестовый контроль</i>	3
ВСЕГО:			16 часов

8. Лабораторный практикум: не предусмотрен

9. Учебно-методическое обеспечение для самостоятельной работы обучающегося

9.1. Содержание самостоятельной работы

№ раздела	Раздел дисциплины	Наименование работ	Трудоемкость (час)
1.	Медицинская статистика	Работа с лекционным материалом, предусматривающая проработку конспекта лекций и учебной литературы	12
		Изучение материала, вынесенного на самостоятельную проработку (отдельные темы, параграфы)	40
		Подготовка к практическому занятию	10
4.	Подготовка к зачету	Повторение и закрепление изученного материала (работа с лекционным материалом, учебной литературой); формулировка вопросов; предэкзаменационные индивидуальные и групповые консультации с преподавателем.	20
ВСЕГО:			82

9.2. Методические указания для обучающихся по освоению дисциплины

Данный раздел программы разработан в качестве самостоятельного документа «Методические рекомендации для аспиранта» в составе УМКД.

10. Ресурсное обеспечение

10.1. Основная литература:

№ п/п	Наименование	Автор(ы)	Год, место издания	Кол-во экземпляров	
				в библиотеке	на кафедре
1	2	3	4	5	6
1.	Медицинская информатика: практикум (2-е изд.).	Гельман В.Я	СПб.: Питер, 2002.	7	5
2.	Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA.	Реброва О.Ю	М.: МедиаСфера, 2002. – 312 с.	5	3
3.	Статистические методы в медико-биологических исследованиях с использованием Excel	Лапач С.Н., Чубренко А.В., Бабич П.Н.	К.: МАРИОН, 2001. – 408 с.	7	2
4.	Руководство по статистике здоровья и здравоохранения	Медик В.А., Токмачев М.С.	М.: ОАО «Издательство «Медицина», 2006.-528 с.	4	2
5.	Медицинская статистика	Лукьянова Е.А.	М.:Изд-во РУДН. 2002 .- 255 с.	7	2
6.	Алгоритм решения задач по медицинской статистике: учеб.-метод. пособие / под ред. В.Н.Савельева.	Савельева В.Ню	Ижевск: ИГМА, 2010.		1

7.	Медицинская статистика: учеб. пособие.	Герасимов А. Н.	М.: МИА, 2007.	1	2
8.	Медицинская информатика: Учебное пособие.	Гусев С.Д.	Красноярск: Издательства, ООО" Версо ", 2009.- 464 с.		2
9.	Прикладная медицинская статистика: учеб. пособие.	Зайцев В. М.	СПб.: Фолиант, 2006.	1	1
10.	Прикладная медицинская статистика: учеб. пособие.	Зайцев В. М.	СПб.: Фолиант, 2003	1	1
11.	.Использование табличного процессора MS Excel для статистической обработки информации о здоровье населения и деятельности медицинской организации Учебно-методическое пособие.	Попова Н.М., Сабитова Н.Г., Толмачев Д.А., Дежина Л.В	Ижевск, 2013 г. – 158 с.		1
12.	Учебное пособие по медицинской статистике: учеб. пособие /под ред. Е. Я. Белицкой .	Белицкая Е.Я.	Л.: Медицина, 1972.		1
13.	Statistica 6. Математическая статистика с элементами теории вероятностей.	Халафян А.А.	М.: Бином, 2010. – 496 с.	1	1
14.	Статистические методы в медико-биологических исследованиях с использованием Excel.	С.Н.Лапач, А.В.Чубенко, П.А.Бабич.	Киев, изд. «Морион», 2000г.		1
15.	Основы организационно-методической службы и статистического анализа в здравоохранении	Э. А. Вальчук, Н. И. Гулицкая, Ф. П. Царук.	Минск: Харвест, 2007. – 400 с.		1

10.2. Дополнительная литература

п/№	Наименование	Автор (ы)	Год, место издания	Кол-во экземпляров	
				в библиотеке	на кафедре
1	2	3	4	5	6
1.	Математическая статистика с элементами теории вероятностей	Халафян А.А.	М.: Издательство Бином, 2011.- 496 с.	3	2
2.	Теория статистики	Громько Г.Л.	М.:ИНФА-М,2009.-476 с.	4	3
3.	Теория вероятностей и математическая статистика	Гурман В.Е.	М.:Издательств о Юрайт,	2	4

			2013.-479 с.		
4.	Медико-биологическая статистика	Гланц С.	М.: Практика, 1998. – 458 с.	1	2
5.	STATISTICA: искусство анализа данных на компьютере	Боровиков В.П	СПб: Питер 2001.-415	2	3
6.	Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере: Учеб. пособие	Боровиков В.П, Ивченко Г.И.	М.: Финансы и статистика, 2000.— 384 с.	1	3
7.	Информационные технологии в медико-биологических исследованиях	Дюк В., Эммануэль В..	СПб.: Питер, 2003.- 321	1	2
8.	Прикладная медицинская статистика	Зайцев В.М., Лифляндский В.Г., Маринкин В.И..	С.-Пб.: Фолиант, 2003.-235 с.	1	2
9.	Телемедицина	Блажис А. К., Дюк. В. А.	Спб.: Спецлит 2001 – 137с.	-	2
10.	Медицинская информатика: практикум.	Гельман В.Я.	СПб: Питер, 2001. -480с. – (Серия "Национальная медицинская библиотека").	-	2
11.	Медицинская статистика: учеб. пособие.	Герасимов А. Н.	М.: МИА, 2007.	-	1
12.	Медицинская информатика: Учебное пособие.	Гусев С.Д.	Красноярск: Издательства, ООО" Версо ", 2009.- 464 с.	1	1
13.	Информационные основы принятия решений в медицине: Учебное пособие.	Карась С.И.	Томск: Печатная мануфактура, 2003.- 145с.	-	1
14.	Информационные основы принятия решений в медицине: Учебное пособие.	Карась С.И.	Томск: Печатная мануфактура, 2003.- 145с.	-	2
15.	Медицинская информатика: Учебник.	Кобринский Б.А., Зарубина Т.В.	М: изд. "Академия", 2009.	1	1
16.	Основы организации автоматизированной обработки персональных данных в медицинских учреждениях. Учебно-методическое пособие.	Кузнецов П.П., Столбов А.П.	М. - 2010. – 38с.	-	1

17.	Медицинские информационные системы: теория и практика/ Под ред. Г.И. Назаренко, Г.С. Осипова.	Назаренко Г.И., Гулиев Я.И., Ермаков Д.Е.	М.: ФИЗМАТЛИТ, 2005. – 320с.	-	1
18.	Общественное здоровье и здравоохранение: учебник под ред. В.А. Миняева, Н.И. Вишнякова –.	Миняева В.А., Вишнякова Н.И.	М., 2004- 144 с.	-	1
19.	Руководство к практическим занятиям по общественному здоровью и здравоохранению, медицинской информатике	В.Н. Савельев, Н.Н. Ежова, В.К. Гасников и др.	Ижевск, 2008. – 432 с.	-	1
20.	Практикум по медицинской информатике: автоматизированное рабочее место врача и системы поддержки принятия врачебного решения.	Санников А.Г., Егоров Д.Б., Скудных А.С., Рухлова С.А.	Тюмень: П.П.Ш., 2009. – 116с.	-	1
21.	Практикум по медицинской информатике: автоматизированное рабочее место врача и системы поддержки принятия врачебного решения.	Санников А.Г., Егоров Д.Б., Скудных А.С., Рухлова С.А.	Тюмень: П.П.Ш., 2009. – 116с	1	1
22.	Учетно-отчетная документация в учреждениях здравоохранения: учеб. пособие	Т.В.Виноградова, О.П.Попова; под ред. В.Н.Савельева	Ижевск: Б.и., 2010	1	1
23.	Медицинская информатика: Учебное пособие.-	Чернов В.И., Родионов О.В., Есауленко И.Э. и др	Воронеж, 2004. – 282с.: ил.	-	1
24.	Экономика здравоохранения: учебное пособие	Решетникова А.В.	М., 2007. – 310 с.	-	1

10.3. Программное обеспечение

При проведении различных видов занятий используются общесистемное и прикладное программное обеспечение, в том числе:

№	Наименование программы	Описание программы
---	------------------------	--------------------

п/п		
1.	Microsoft EXCEL	Встроенные статистические функции. Пакет анализа данных.
2.	Statistica for Windows	Мощный статистический пакет для обработки данных.
3.	SPSS	Мощный статистический пакет для обработки данных.

10.4. Ресурсы информационно-телекоммуникативной сети «Интернет»

Базы данных, информационно-поисковые системы:

№ п/п	Интернет ресурс (адрес)	Описание ресурса
1.	http://www.nsu.ru/mmftvims/chernova/tv/	Чернова Н.И., НГУ. Семестровый курс лекций по теории вероятностей для студентов экономического факультета
2.	http://www.statsoft.ru/home/textbook/default.htm	Электронный учебник по прикладной статистике фирмы StatSoft
3.	http://teorver-online.narod.ru/	Интернет-учебник «Теория вероятностей и математическая статистика» для студентов естественных факультетов
4.	http://www.ksu.ru/infres/volodin/	Володин И.Н., Казанский ГУ. Лекции по теории вероятностей и математической статистике
5.	http://www.tvp.ru/vnizd/mathem4.htm	Эконометрика. Временные ряды. [Электронный ресурс].
6.	http://crow.academy.ru/econometrics/	Прикладная эконометрика. МГУ им Ломоносова [Электронный ресурс].
7.	www.cemi.rssi.ru	Центральный экономико-математический институт РАН (цэми) [Электронный ресурс].

10.5. Материально-техническое обеспечение дисциплины

Лекционные занятия:

- - комплект электронных презентаций по темам лекций;
- - аудитория, оснащенная презентационной техникой (мультимедийным проектором, экран, компьютер/ноутбук).

Практические занятия:

Аспиранту предоставляется возможность использования учебных аудиторий, оснащенных наглядными пособиями.

Обеспечивается доступом к персональному компьютеру с выходом в интернет и доступам к научным базам данных.

Предоставляется возможность использования:

- Научной литературы на кафедре и библиотеке университета.
- Справочными пособиями и справочными таблицами.

- Комплектами тренировочных тестовых заданий
- Комплектами образцов выполнения практических заданий по дисциплине «Математическая статистика»

11. Использование инновационных (активных и интерактивных) методов обучения

Используемые активные методы обучения при изучении данной дисциплины составляют _____% от объема аудиторных занятий.

№	Наименование раздела (перечислить те разделы, в которых используются активные и/или интерактивные формы (методы) обучения)	Формы занятий с использованием активных и интерактивных методов обучения	Трудоемкость (час.)
1	<i>Медицинская статистика</i>	<i>Лекция 1. «Различные определения вероятности. Теоремы сложения и умножения вероятностей. Теоремы сложения и умножения вероятностей. Повторные независимые испытания». Лекция-визуализация.</i>	2
		<i>Лекция 2 «Случайные величины и их характеристики. Свойства математического ожидания и дисперсии» Лекция-визуализация</i>	2

12. Фонд оценочных средств для текущего контроля

Примеры оценочных средств текущего контроля:

Инструкция по выполнению тестовых заданий:

Из предложенных ответов для каждого теста, необходимо выбрать только один правильный ответ.

Тест 1. Число способов, которым можно выбрать двух человек из трех равно ...:

- А.1
- Б.2
- В.3
- Г.4

Тест 2. Число трехбуквенных слов из букв слова «ромб» равно ...

- А.2
- Б.3

В.4

Г.5

Тест 3. Вероятность попадания при одном выстреле 0,9, тогда вероятность трех промахов при трех выстрелах равна ...

А. 0,001

Б. 0,5

В. 0,01

Г. 0,005

Тест 4. Вероятность угадывания последней цифры телефонного номера ровно с двух раз равна ...

А. 0,2

Б. 0,1

В. 0,3

Г. 0,5

Тест 5. Число различных очередей из трех человек равно ...

А. 3

Б. 4

В. 6

Г. 8

Тест 6. Элементарное событие – это ...

А. эксперимент

Б. число

В. исход эксперимента

Г. вывод

Тест 7. Событие – это ...

А. утверждение

Б. подмножество

В. пространство элементарных событий

Г. доказательство

Тест 8. Вероятность – это ...

А. функция на пространстве элементарных событий

Б. утверждение

В. множество

Г. эксперимент

Тест 9. $P(A+B)=...$

А. $P(A)+P(B)-P(AB)$

Б. $P(A)-P(B)$

В. $P(AB)+P(A)$

Г. $P(AB)+P(B)$

Тест 10. Случайная величина – это ...

А. доказанное утверждение

Б. измеримая функция

В. очевидное свойство

Г. положительное число

Эталон правильных ответов на тестовые задания

№ п/п	Ответы на тестовые задания				
	Тест 1	Тест 2	Тест 3	Тест 4	Тест 5
Тест	В	В	А	Б	В
Тест	Тест 6	Тест 7	Тест 8	Тест 9	Тест 10
Ответы	В	Б	А	А	Б

Критерии оценки текущего контроля:

оценка «отлично» выставляется студенту, если тест решен на 90% и более;

оценка «хорошо» выставляется студенту, если тест решен на 80-90%;

оценка «удовлетворительно» выставляется студенту, если тест решен на 60-80%;

оценка «неудовлетворительно» выставляется студенту, если тест решен на 60%.

12. Фонд оценочных средств для промежуточной аттестации (зачет)

Процедура проведения промежуточной аттестации (зачета)

Допуск аспиранта к сдаче зачета по дисциплине «Медицинская статистика» возможен при условии отсутствия у аспиранта академических задолженностей по дисциплине.

Зачет проводится в форме устного собеседования по вопросам билета. Билет включает 2 теоретических вопроса и одну задачу. Ответы на вопросы аспирант должен сопровождать конкретными примерами и ссылками на реальные обстоятельства и ситуации; при этом высказать свою точку зрения по излагаемым вопросам.

На подготовку к ответу дается 45 минут, в течение которых аспирант записывает тезисы ответов на специальных листах, выдаваемых вместе с билетом. Тезисы должны быть записаны понятным почерком.

Преподаватель, принимающий зачет по дисциплине «Медицинская статистика» имеет право задавать дополнительные вопросы по билету для уточнения степени знаний аспиранта. Преподаватель выставляет оценку аспиранту по каждому вопросу билета и каждому дополнительному вопросу.

Оценки объявляются аспиранту в день сдачи зачета.

Примеры оценочных средств промежуточной аттестации по дисциплине «Медицинская статистика»

Вопросы для сдачи зачета по дисциплине «Медицинская статистика»:

1. Основные понятия теории вероятностей: испытание, событие, виды событий, примеры.
2. Классическое определение вероятности события.
3. Теорема сложения вероятностей совместных и несовместных событий.
4. Теорема умножения вероятностей зависимых и независимых событий. Условная вероятность.
5. Полная группа событий. Противоположные события.
6. Формула полной вероятности.
7. Повторение независимых испытаний. Формула Бернулли, формула Пуассона локальная и интегральная формулы Муавра-Лапласа. Наивероятнейшее число наступления события.
8. Дискретная случайная величина, закон ее распределения. Основные числовые характеристики дискретной случайной величины.
9. Непрерывная случайная величина. Основные числовые характеристики непрерывной случайной величины.
10. Нормальный закон распределения непрерывной случайной величины.
11. Понятие о законе больших чисел.
12. Вариационный ряд. Виды вариационных рядов их графическое изображение.
13. Числовые характеристики вариационного ряда.
14. Генеральная и выборочная совокупности.
15. Выборка: виды, способы образования. Основная задача выборочного метода.
16. Понятие об интервальном оценивании. Доверительная вероятность, доверительный интервал.
17. Статистическая гипотеза, статистический критерий.
18. Уровень значимости и мощность критерия.
19. Построение теоретического закона распределения по опытным данным.
20. Понятие о критериях согласия.
21. Критерий Пирсона χ^2 и схема его применения.
22. Функциональная, статистическая и корреляционная зависимости.
23. Основные задачи теории корреляции.
24. Линейная регрессия. Уравнения регрессии.
25. Коэффициент корреляции: оценка тесноты и вида связи между признаками X и Y.

Задачи для промежуточной аттестации

1. С целью определения средней суммы кредитов в банке, имеющем 2000 клиентов, по схеме собственно-случайной бесповторной выборки проведено обследование 100 кредитов. Результаты обследования представлены в таблице:

Сумма кредита, тыс. руб.	50 - 150	150 - 250	250 - 350	350 - 450	450 - 550	Итого
Число клиентов	14	24	35	20	7	100

Найти: а) границы, в которых с вероятностью 0,9488 находится средняя сумма всех кредитов в банке; б) объем бесповторной выборки, при котором те же границы для средней суммы кредитов в банке (см. п. а)) можно гарантировать с вероятностью 0,9; в) вероятность того, что доля всех клиентов, у которых сумма кредита больше 250 тыс. руб., отличается от доли таких клиентов в выборке не более чем на 0,1 (по абсолютной величине).

2. По данным задачи 1, используя критерий χ^2 - Пирсона, при уровне значимости $\alpha = 0,05$ проверить гипотезу о том, что случайная величина X – сумма кредита – распределена по нормальному закону. Построить на одном чертеже гистограмму эмпирического распределения и соответствующую нормальную кривую.

3. Распределение 250 пациентов по возрасту мужчин X (лет) и женщин Y (лет) представлено в таблице:

$x \backslash y$	15 - 25	25 - 35	35 - 45	45 - 55	55 - 65	Итого:
15 - 25	7	3				10
25 - 35	52	110	13	1		176
35 - 45	1	14	23	2		40
45 - 55		1	4	6	1	12
55 - 65				3	6	9
65 - 75					3	3
Итого:	60	128	40	12	10	250

Необходимо:

1) Вычислить групповые средние \bar{x}_j и \bar{y}_i , построить эмпирические линии регрессии.

2) Предполагая, что между переменными X и Y существует линейная корреляционная зависимость: а) найти уравнения прямых регрессии, построить их графики на одном чертеже с эмпирическими линиями регрессии и дать содержательную интерпретацию полученных уравнений; б) вычислить коэффициент корреляции на уровне значимости $\alpha = 0,05$, оценить его значимость и сделать вывод о тесноте и направлении связи между переменными X и Y .

Эталон решения задач промежуточной аттестации аспирантов

1. Решение.

От интервального распределения перейдем к дискретному, взяв в качестве представителя интервала его середину \tilde{x}_i .

Для расчета выборочной средней и выборочной дисперсии составим таблицу.

Сумма вклада, тыс. руб.	Количество вкладов, n_i	Середина, x_i	$x_i - C$	$\frac{x_i - C}{k}$	$\frac{x_i - C}{k} \cdot n_i$	$\left(\frac{x_i - C}{k}\right)^2 \cdot n_i$
50 - 150	14	100	-200	-2	-28	56
150 - 250	24	200	-100	-1	-24	24
250 - 350	35	300	0	0	0	0
350 - 450	20	400	100	1	20	20
450 - 550	7	500	200	2	14	28
Суммы	100				-18	128

$C = 300$ - середина интервала с наибольшей частотой;

$k = 100$ - величина интервала.

Выборочное среднее найдем по формуле $\bar{x} = \frac{\sum_{i=1}^m \frac{x_i - C}{k} \cdot n_i}{n} \cdot k + C$

$$\bar{x}_a = \frac{-18}{100} \cdot 100 + 300 = 282 \text{ тыс. руб.}$$

Выборочная дисперсия

$$\sigma^2 = \frac{\sum_{i=1}^m \left(\frac{x_i - C}{k}\right)^2 \cdot n_i}{n} \cdot k^2 - (\bar{x} - C)^2,$$

$$\sigma^2 = \frac{128}{100} \cdot 100^2 - (282 - 300)^2 = 12476.$$

Выборочное среднее квадратическое отклонение

$$\bar{\sigma}_a = \sqrt{\sigma_a^2} = \sqrt{12476} \approx 111,696.$$

а) Средняя квадратическая ошибка среднего значения признака для

бесповторной выборки $\sigma_{\bar{x}_a} = \sqrt{\frac{\bar{\sigma}_a^2}{n} \cdot \left(1 - \frac{n}{N}\right)}$.

Число всех вкладов $N = 2000$, объем выборки $n = 100$

$$\sigma_{\bar{x}_a} = \sqrt{\frac{12476}{100} \cdot \left(1 - \frac{100}{2000}\right)} \approx \sqrt{118,5520} \approx 10,8868.$$

Вероятности $\beta = 0,9488$ соответствует $t = 1,95$, так как $\Phi(1,95) = 0,9488$.

Предельная ошибка $\Delta = t \cdot \sigma_{\bar{x}_a} = 1,95 \cdot 10,8868 \approx 21,2270$.

Нижняя граница $\bar{x}_a - \Delta = 282 - 21,227 = 260,773$,

верхняя граница $\bar{x}_a + \Delta = 282 + 21,227 = 303,227$.

С вероятностью 0,9488 средняя сумма всех вкладов в сберегательном банке заключена в границах от 260,773 до 303,227 тыс. руб.

б) Вероятности $P = 0,9$ соответствует $t = 1,64$, так как $\Phi(1,64) = 0,9$.

Число клиентов, которых надо обследовать для повторной выборки

$$n_x = \frac{t^2 \sigma_a^2}{\Delta^2} = \frac{1,64^2 \cdot 12476}{21,227^2} \approx 74,912.$$

Для бесповторной выборки

$$n'_x = \frac{n_x \cdot N}{n_x + N} = \frac{74,912 \cdot 2000}{74,912 + 2000} \approx 72,207. \text{ Округляем до большего целого } 73.$$

Чтобы с вероятностью 0,9 гарантировать те же границы для средней суммы всех кредитов в банке, что и в п. а) объем бесповторной выборки должен быть равным 73 кредитам.

в) Выборочная доля клиентов, у которых сумма кредита больше 250 тыс. руб., равна $\omega = \frac{35 + 20 + 7}{100} = 0,62$.

Средняя квадратическая ошибка доли для бесповторной выборки

$$\sigma'_\omega = \sqrt{\frac{\omega(1-\omega)}{n} \cdot \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,62 \cdot (1-0,62)}{100} \cdot \left(1 - \frac{100}{2000}\right)} \approx \sqrt{0,002238} \approx 0,0473 \approx 0,047.$$

Предельная ошибка $\Delta = 0,1$. $t_\beta = \frac{\Delta}{\sigma_{\bar{x}_e}} = 0,1 / 0,0473 \approx 2,11$.

Находим требуемую вероятность $P = \Phi(t_\beta) = \Phi(2,11) = 0,9651$

Вероятность того, что доля всех клиентов, у которых сумма кредита больше 250 тыс. руб., отличается от доли таких клиентов в выборке не более чем на 0,1 (по абсолютной величине), приближенно равна 0,9651.

2. Решение.

Проверяется гипотеза H_0 : случайная величина X – сумма вклада – распределена по нормальному закону. Функция плотности вероятности и функция распределения имеют вид

$$\varphi_n(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad F_n(x) = \frac{1}{2} + \frac{1}{2} \cdot \Phi\left(\frac{x-a}{\sigma}\right), \text{ где } a, \sigma - \text{параметры распределения.}$$

В качестве оценок этих параметров возьмем выборочное среднее значение и дисперсию.

$$a \approx \bar{x} = 282; \quad \sigma = \bar{\sigma}_a = \sqrt{\bar{\sigma}_a^2} = \sqrt{12476} \approx 111,696.$$

$$\text{Тогда } \varphi_i(x) = \frac{1}{111,696\sqrt{2\pi}} e^{-\frac{(x-282)^2}{2 \cdot 12476}} \text{ и } F_i(x) = \frac{1}{2} + \frac{1}{2} \cdot \Phi\left(\frac{x-282}{111,696}\right).$$

Вычислим наблюдаемое значение критерия Пирсона по формуле

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - np_i)^2}{np_i}, \text{ где}$$

m - число интервалов; n_i - частота (эмпирическая); n - объем выборки; p_i - теоретическая

вероятность попадания случайной величины в i -ый интервал; np_i - теоретическая частота.

Вероятность p_i попадания случайной величины X в интервал $(x_i; x_{i+1})$ найдем по формуле

$$p_i = P(x_i < X < x_{i+1}) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{x_{i+1} - a}{\sigma} \right) - \hat{\Phi} \left(\frac{x_i - a}{\sigma} \right) \right] = \frac{1}{2} \left[\hat{\Phi} \left(\frac{x_{i+1} - 282}{111,696} \right) - \hat{\Phi} \left(\frac{x_i - 282}{111,696} \right) \right].$$

$$p_1 = P(50 < X < 150) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{150 - 282}{111,696} \right) - \hat{\Phi} \left(\frac{50 - 282}{111,696} \right) \right] =$$

$$= 0,5 \cdot (\Phi(-1,18) - \Phi(-2,08)) = 0,5 \cdot (-0,7620 + 0,9625) = 0,1002.$$

$$p_2 = P(150 < X < 250) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{250 - 282}{111,696} \right) - \hat{\Phi} \left(\frac{150 - 282}{111,696} \right) \right] =$$

$$= 0,5 \cdot (\Phi(-0,29) - \Phi(-1,18)) = 0,5 \cdot (-0,2282 + 0,7620) = 0,2669.$$

$$p_3 = P(250 < X < 350) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{350 - 282}{111,696} \right) - \hat{\Phi} \left(\frac{250 - 282}{111,696} \right) \right] =$$

$$= 0,5 \cdot (\Phi(0,61) - \Phi(-0,29)) = 0,5 \cdot (0,4581 + 0,2282) = 0,3432.$$

$$p_4 = P(350 < X < 450) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{450 - 282}{111,696} \right) - \hat{\Phi} \left(\frac{350 - 282}{111,696} \right) \right] =$$

$$= 0,5 \cdot (\Phi(1,50) - \Phi(0,61)) = 0,5 \cdot (0,8664 - 0,4581) = 0,2041.$$

$$p_5 = P(450 < X < 550) = \frac{1}{2} \left[\hat{\Phi} \left(\frac{550 - 282}{111,696} \right) - \hat{\Phi} \left(\frac{450 - 282}{111,696} \right) \right] =$$

$$= 0,5 \cdot (\Phi(2,40) - \Phi(1,50)) = 0,5 \cdot (0,9836 - 0,8664) = 0,0586.$$

Для расчета составим вспомогательную таблицу

i	Интервал $(X_i; X_{i+1})$	Эмпирические частоты n_i	Вероятность p_i	Теоретические частоты np_i	$n_i - np_i$	$(n_i - np_i)^2$	$(n_i - np_i)^2 / np_i$
1	50 - 150	14	0,1002	10,020	3,980	15,8404	1,5809
2	150 - 250	24	0,2669	26,690	-2,690	7,2361	0,2711
3	250 - 350	35	0,3432	34,320	0,680	0,4624	0,0135
4	350 - 450	20	0,2041	20,410	-0,410	0,1681	0,0082
5	450 - 550	7	0,0586	5,860	1,140	1,2996	0,2218
	Суммы	100	0,9730	97,300			2,0955

$$\chi_{набл}^2 = 2,0955.$$

Найдем по таблице критическое значение критерия $\chi_{кр}^2 = \chi_{\alpha, k}^2$, $k = m - s - 1$, $m = 5$ - число интервалов, $s = 2$ - число параметров распределения, $\alpha = 0,05$ - уровень значимости, $k = 5 - 2 - 1 = 2$, $\chi_{\alpha}^2 = \chi_{0,05; 2}^2 = 5,99$.

Сравниваем наблюдаемое значение критерия с критическим $\chi_{набл}^2 < \chi_{кр}^2$

$2,0955 < 5,99$. Это означает, что наблюдаемое значение не попало в критическую область. Поэтому гипотеза о нормальном распределении размера кредита согласуется с данными выборки и должна быть принята.

Гистограмма - это совокупность прямоугольников, основаниями которых служат частичные интервалы $(x_i; x_{i+1}]$, а высота которых равна $\omega_i = \frac{n_i}{n \cdot k_i}$.

$k_i = x_{i+1} - x_i$ - длина частичного интервала, $k_i = 100$, $n \cdot k_i = 100 \cdot 100 = 10000$

$$\omega_1 = \frac{14}{10000} = 0,0014, \quad \omega_2 = \frac{24}{10000} = 0,0024, \quad \omega_3 = \frac{35}{10000} = 0,0035, \quad \omega_4 = \frac{20}{10000} = 0,0020,$$

$$\omega_5 = \frac{7}{10000} = 0,0007.$$

Для построения графика нормальной кривой отметим точки $(x_i; p_i/k)$, где x_i - середина интервала, p_i - вероятность попадания в интервал.

Вершина при $x = a = 282$.

$$y_{\max} = \frac{0,3989}{\sigma} = \frac{0,3989}{111,696} = 0,003574.$$

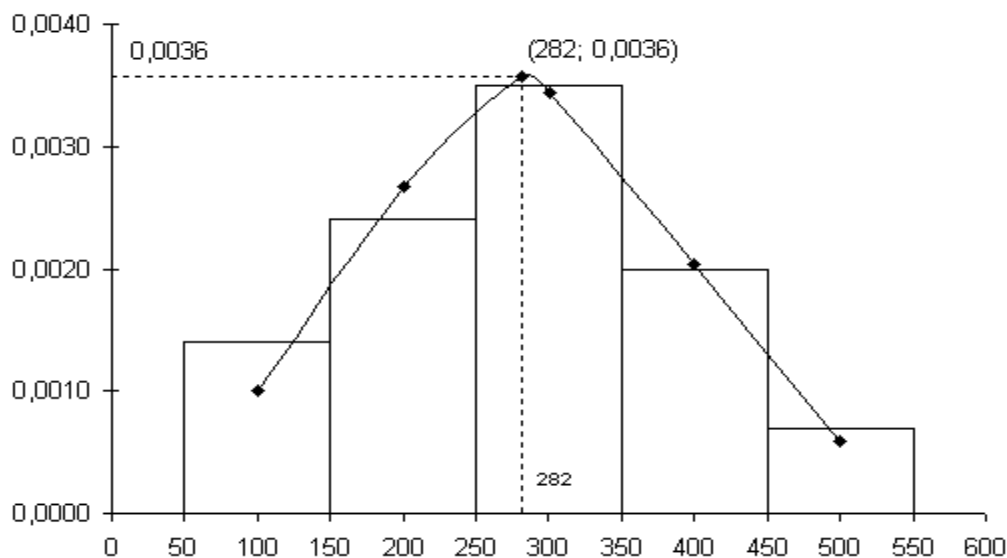
$$p_1 / k = 0,1002 / 100 = 0,0010$$

$$p_2 / k = 0,2669 / 100 = 0,0027$$

$$p_3 / k = 0,3432 / 100 = 0,0034$$

$$p_4 / k = 0,2041 / 100 = 0,0020$$

$$p_5 / k = 0,0586 / 100 = 0,0006$$



3. Решение.

По исходным данным составим корреляционную таблицу, где интервалы представлены своими серединами.

$x_i \backslash y_j$	20	30	40	50	60	n_i
20	7	3				10

30	52	110	13	1		176
40	1	14	23	2		40
50		1	4	6	1	12
60				3	6	9
70					3	3
n_j	60	128	40	12	10	250

1) Найдем групповые средние по Y по формуле $\bar{y}_{x_i} = \frac{\sum_{j=1}^t y_j n_{ij}}{n_{x_i}}$.

$$x_1 = 20 \quad \bar{y}_1 = \bar{y}_{x_1=20} = (20 \cdot 7 + 30 \cdot 3) / 10 = 230 / 10 = 23,000$$

$$x_2 = 30 \quad \bar{y}_2 = \bar{y}_{x_2=30} = (20 \cdot 52 + 30 \cdot 110 + 40 \cdot 13 + 50 \cdot 1) / 176 = 4910 / 176 = 27,898$$

$$x_3 = 40 \quad \bar{y}_3 = \bar{y}_{x_3=40} = (20 \cdot 1 + 30 \cdot 14 + 40 \cdot 23 + 50 \cdot 2) / 40 = 1460 / 40 = 36,500$$

$$x_4 = 50 \quad \bar{y}_4 = \bar{y}_{x_4=50} = (30 \cdot 1 + 40 \cdot 4 + 50 \cdot 6 + 60 \cdot 1) / 12 = 550 / 12 = 45,833$$

$$x_5 = 60 \quad \bar{y}_5 = \bar{y}_{x_5=60} = (50 \cdot 3 + 60 \cdot 6) / 9 = 510 / 9 = 56,667$$

$$x_6 = 70 \quad \bar{y}_6 = \bar{y}_{x_6=70} = 60 \cdot 3 / 3 = 60,000$$

Составим таблицу 2.

Таблица 2

x_i	20	30	40	50	60	70
\bar{y}_i	23,000	27,898	36,500	45,833	56,667	60,000

По точкам $(x_i; \bar{y}_i)$ построим эмпирическую линию регрессии Y на X. Эти точки расположены вблизи прямой с уравнением $y = ax + b$, где a и b неизвестные параметры и их нужно определить.

Групповые средние по X найдем по формуле $\bar{x}_j = \frac{\sum_{i=1}^m x_i n_{ij}}{n_j}$.

$$y_1 = 20 \quad \bar{x}_1 = \bar{x}_{y_1=20} = (20 \cdot 7 + 30 \cdot 52 + 40 \cdot 1) / 60 = 1740 / 60 = 29,000$$

$$y_2 = 30 \quad \bar{x}_2 = \bar{x}_{y_2=30} = (20 \cdot 3 + 30 \cdot 110 + 40 \cdot 14 + 50 \cdot 1) / 128 = 3970 / 128 = 31,016$$

$$y_3 = 40 \quad \bar{x}_3 = \bar{x}_{y_3=40} = (30 \cdot 13 + 40 \cdot 23 + 50 \cdot 4) / 40 = 1510 / 40 = 37,750$$

$$y_4 = 50 \quad \bar{x}_4 = \bar{x}_{y_4=50} = (30 \cdot 1 + 40 \cdot 2 + 50 \cdot 6 + 60 \cdot 3) / 12 = 590 / 12 = 49,167$$

$$y_5 = 60 \quad \bar{x}_5 = \bar{x}_{y_5=60} = (50 \cdot 1 + 60 \cdot 6 + 70 \cdot 3) / 10 = 620 / 10 = 62,000$$

Составим таблицу 3

Таблица 3

\bar{x}_j	29,000	31,016	37,750	49,167	62,000
y_j	20	30	40	50	60

По точкам $(\bar{x}_j; y_j)$ построим эмпирическую линию регрессии X на Y. Эти точки расположены вблизи прямой с уравнением $x = cy + d$, где c и d неизвестные параметры и их нужно определить.

Для получения уравнений прямых регрессий вычислим выборочные средние

$$\bar{x} = \frac{\sum_{i=1}^m x_i n_{x_i}}{n} \quad \text{и} \quad \bar{y} = \frac{\sum_{j=1}^t y_j n_{y_j}}{n}.$$

$$\bar{x} = \frac{20 \cdot 10 + 30 \cdot 176 + 40 \cdot 40 + 50 \cdot 12 + 60 \cdot 9 + 70 \cdot 3}{250} = \frac{8430}{250} = 33,72$$

$$\bar{y} = \frac{20 \cdot 60 + 30 \cdot 128 + 40 \cdot 40 + 50 \cdot 12 + 60 \cdot 10}{250} = \frac{7840}{250} = 31,36$$

Выборочные дисперсии находим по формулам $\sigma_x^2 = \overline{x^2} - \bar{x}^2$ и $\sigma_y^2 = \overline{y^2} - \bar{y}^2$

$$\overline{x^2} = \frac{20^2 \cdot 10 + 30^2 \cdot 176 + 40^2 \cdot 40 + 50^2 \cdot 12 + 60^2 \cdot 9 + 70^2 \cdot 3}{250} = \frac{303500}{250} = 1214$$

$$\sigma_x^2 = 1214 - 33,72^2 = 76,9616.$$

$$\overline{y^2} = \frac{20^2 \cdot 60 + 30^2 \cdot 128 + 40^2 \cdot 40 + 50^2 \cdot 12 + 60^2 \cdot 10}{250} = \frac{269200}{250} = 1076,8$$

$$\sigma_y^2 = 1076,8 - 31,36^2 = 93,3504.$$

Вычислим средние квадратические отклонения

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{76,9616} \approx 8,7728; \quad \sigma_y = \sqrt{\sigma_y^2} = \sqrt{93,3504} \approx 9,6618.$$

Вычислим $\mu = \overline{xy} - \bar{x} \cdot \bar{y}$ по формуле $\mu = \frac{\sum_{i=1}^m \sum_{j=1}^t x_i \cdot y_j \cdot n_{ij}}{n} - \bar{x} \cdot \bar{y}$.

$$\begin{aligned} \mu &= (20 \cdot 20 \cdot 7 + 20 \cdot 30 \cdot 3 + 30 \cdot 20 \cdot 52 + 30 \cdot 30 \cdot 110 + 30 \cdot 40 \cdot 13 + 30 \cdot 50 \cdot 1 + \\ &+ 40 \cdot 20 \cdot 1 + 40 \cdot 30 \cdot 14 + 40 \cdot 40 \cdot 23 + 40 \cdot 50 \cdot 2 + 50 \cdot 30 \cdot 1 + 50 \cdot 40 \cdot 4 + \\ &+ 50 \cdot 50 \cdot 6 + 50 \cdot 60 \cdot 1 + 60 \cdot 50 \cdot 3 + 60 \cdot 60 \cdot 6 + 70 \cdot 60 \cdot 3) / 250 - 33,72 \cdot 31,36 = \\ &= 281000 / 250 - 1057,4592 = 1124 - 1057,4592 = 66,5408. \end{aligned}$$

Вычислим коэффициенты регрессии по формулам

$$\rho_{y/x} = \frac{\mu}{\sigma_x^2} = 66,5408 : 76,9616 \approx 0,8646 \approx 0,865;$$

$$\rho_{x/y} = \frac{\mu}{\sigma_y^2} = 66,5408 : 93,3504 \approx 0,7128 \approx 0,713.$$

а) Составим уравнение регрессии X на Y $x - \bar{x} = \rho_{x/y} (y - \bar{y})$

$$x - 33,72 = 0,713 \cdot (y - 31,36) \quad \text{или} \quad x = 0,713 y + 11,366.$$

Прямую проведем через точки (33,72; 31,36) и (11,366; 0,00).

Уравнение регрессии X на Y показывает средний возраст мужчины, вступившего в брак с женщиной возраста y.

Содержательный смысл коэффициента регрессии $\rho_{x/y} = \frac{\mu}{\sigma_y^2} = 0,713$ состоит в том, что при увеличении возраста женщины, вступающей в брак, на 1 год возраст супруга увеличивается в среднем на 0,713 года.

Составим уравнение регрессии Y на X $y - \bar{y} = \rho_{y/x}(x - \bar{x})$

$$y - 31,36 = 0,865 \cdot (x - 33,36) \quad \text{или} \quad y = 0,865x + 2,206.$$

Прямую проведем через точки (33,72; 31,36) и (0,00; 2,206).

Уравнение регрессии Y на X показывает средний возраст женщины, вступившей в брак с мужчиной возраста x.

Содержательный смысл коэффициента регрессии $\rho_{y/x} = \frac{\mu}{\sigma_x^2} = 0,865$ состоит в том, что при увеличении возраста мужчины, вступающего в брак, на 1 год возраст супруги увеличивается в среднем на 0,865 года.

б) Коэффициент корреляции $r = \frac{\mu}{\sigma_x \sigma_y} = \frac{66,5408}{8,7728 \cdot 9,66181} \approx 0,7850$.

Для проверки значимости коэффициента корреляции вычислим наблюдаемое значение

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}; \quad t_{\text{табл}} = \frac{0,7850 \cdot \sqrt{250-2}}{\sqrt{1-0,7850^2}} \approx 19,958.$$

Критическое значение для уровня значимости $\alpha = 0,05$ и числа степеней свободы $k = n-2 = 250-2 = 248$ находим по таблице $t_{1-0,05;248} = t_{0,95;248} = 1,97$.

Получили $|t_{\text{набл}}| > t_{\text{кр}}$, так как $19,958 > 1,97$.

Следовательно, коэффициент корреляции значимо отличается от нуля.

Коэффициент корреляции $r = 0,7851 > 0$ и попадает по абсолютной величине в интервал 0,7 - 0,99. Следовательно, между возрастом вступающих в брак мужчины (X) и женщины (Y) существует прямая сильная корреляционная связь. При увеличении (уменьшении) значения одной величины соответственно увеличивается (уменьшается) среднее значение другой.

Критерии оценки:

оценка «отлично» выставляется аспиранту, если он без замечаний ответил на все вопросы билета (более 90%) и решил задачу;

оценка «хорошо» выставляется аспиранту, если он без существенных замечаний ответил на все вопросы билета (80-90%), однако затруднился с отдельными вопросами задачи;

оценка «удовлетворительно» выставляется аспиранту, если он ответил кратко и неполно на вопросы билета без грубых ошибок (65-80%), решил задачу не полностью;

оценка «неудовлетворительно» выставляется аспиранту, если он не ответил правильно на вопросы билета (менее 65%), не решил задачу.

--	--	--	--	--