

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе

Белорусского государственного университета

_____ А.В.Данильченко

(подпись)

02.07.2014 г.

(дата утверждения)

Регистрационный № УД-1203/баз.

ПРИКЛАДНАЯ СТАТИСТИКА

**Учебная программа учреждения высшего образования по учебной дисциплине
для специальности:**

1-23 01 15 Социальные коммуникации

2014 г.

СОСТАВИТЕЛЬ:

О. В. Терещенко, заведующая кафедрой социальной коммуникации Белорусского государственного университета, кандидат социологических наук, доцент

РЕЦЕНЗЕНТЫ:

Е. А. Кечина, профессор кафедры социологии Белорусского государственного университета, доктор социологических наук, профессор

З. В. Пунчик, доцент кафедры экономической информатики Учреждения образования «Белорусский государственный экономический университет», кандидат социологических наук, доцент

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой социальной коммуникации факультета философии и социальных наук Белорусского государственного университета (протокол № 11 от 23.05.2014 г.);

Научно-методическим советом Белорусского государственного университета (протокол № 6 от 20.06.2014 г.)

Ответственный за редакцию: Терещенко О. В.

Ответственный за выпуск: Терещенко О. В.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебная дисциплина «Прикладная статистика» предназначена для реализации на первой ступени высшего образования.

Учебная программа учреждения высшего образования по учебной дисциплине «Прикладная статистика» разработана в соответствии с Образовательным стандартом и учебным планом по специальности 1-23 01 15 Социальные коммуникации.

Цель учебной дисциплины – обучение студентов методам прикладной статистики, предназначенным для профессиональной обработки и анализа данных количественных социальных исследований, и их практическому применению, в том числе с использованием компьютерного программного обеспечения.

Задачи учебной дисциплины:

- рассмотреть особенности статистического подхода к анализу социальной реальности;
- обучить студентов основам измерений, используемых в социальных науках, построению индикаторов и индексов, в том числе относительных показателей;
- обучить студентов построению одномерных распределений, их графическому представлению; вычислению и интерпретации мер центра распределения и разброса данных;
- ознакомить студентов с основами статистического вывода; обучить оцениванию параметров генеральной совокупности и репрезентативности выборки, статистической проверке гипотез;
- рассмотреть основные модели статистической связи между переменными; обучить выбору, применению и интерпретации мер статистической связи в зависимости от используемой модели связи и уровня измерений переменных;
- обучить представлению данных социальных исследований в компьютерных статистических программах и решению задач статистического анализа данных с использованием статистического программного обеспечения.

Дисциплина «Прикладная статистика» читается во втором и третьем семестрах, после дисциплины «Основы высшей математики». Она связана также с дисциплинами «Методология и методы исследований коммуникации» (3–6 семестры) и «Социально-экономическая статистика» (3–4 семестры), прохождением студентами учебной практики (4, 6 семестры), написанием курсовых и дипломной работ.

По завершению изучения дисциплины «Прикладная статистика» студент должен

знать:

- уровни измерений переменных;
- правила группировки данных;
- меры центра распределения и степени разброса данных;
- основы статистического вывода;
- модели и меры статистических связей;

уметь:

- измерять переменные в социальных исследованиях;
- строить одномерные распределения и группировки;
- строить одномерные и двумерные графики;
- оценивать параметры генеральной совокупности и репрезентативность выборки;
- проверять статистические гипотезы;
- выбирать модели, вычислять и интерпретировать меры статистической связи;

владеть:

- навыками постановки статистических задач, формулирования статистических гипотез
- компьютерным программным обеспечением, предназначенным для решения задач статистического анализа данных.

На изучение дисциплины «Прикладная статистика» учебным планом отводится 302 учебных часа, из них 140 – аудиторных, примерное распределение которых по видам занятий включает 70 лекционных, 56 часов практических занятий и 14 часа контролируемой самостоятельной работы.

Рекомендуемая форма отчетности – два экзамена.

ПРИМЕРНЫЙ ТЕМАТИЧЕСКИЙ ПЛАН

№	Наименование раздела, темы	Кол-во аудиторных часов		
		Лекции	Практич. занятия	Ауд. контроль УСР
РАЗДЕЛ 1. Статистика как наука		10	2	2
1	Природа статистики. Особенности статистического подхода в социальных науках	2		
2	Измерения в социальных исследованиях	6	2	
3	Представление данных социальных исследований	2		2
РАЗДЕЛ 2. Дескриптивный анализ данных		20	16	6
4	Одномерные частотные распределения для дискретных и непрерывных переменных	8	6	
5	Графическое представление одномерных распределений	4	4	4
6	Показатели центра распределения и разброса данных	4	4	
7	Анализ формы распределения. Стандартизация переменных	4	2	2
РАЗДЕЛ 3. Основы статистического вывода		12	16	6
8	Теория вероятностей как основа статистического вывода	2		
9	Теоретические распределения, статистические таблицы	2	6	2
10	Оценивание параметров генеральной совокупности. Репрезентативность выборки	4	4	2
11	Статистическая проверка гипотез	4	6	2
РАЗДЕЛ 4. Анализ статистических связей		28	14	8
12	Понятия статистической связи и независимости	2		
13	Частотные модели статистической связи. Таблица сопряженности	8	6	2
14	Меры связи, основанные на рангах	4	2	2
15	Линейная модель парной связи. Корреляционный и регрессионный анализ	4	2	2
16	Нелинейные модели парной связи	4	2	
17	Основы планирования эксперимента. Дисперсионный анализ	6	2	2
ИТОГО		70	56	14

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

РАЗДЕЛ 1. Прикладная статистика в социальных исследованиях**Тема 1.** Природа статистики. Особенности статистического подхода в социальных науках

Статистика как наука, имеющая дело со сбором, обработкой, анализом и интерпретацией данных о массовых явлениях и процессах. Статистические совокупности: генеральные, выборочные. Элементы статистических совокупностей (случаи). Дизайн статистического социального исследования: сплошной и выборочный подходы. Виды исследований генеральной совокупности: количественные, качественные. Количественные (статистические) методы исследования: сплошное обследование, выборочное обследование, факторный эксперимент. Временной подход в социальных исследованиях: мониторинговое, лонгитюдное (когортное) исследования. Функции статистики: описание, обобщение, объяснение (прогнозирование). Особенности статистического подхода в социальных науках: абстрагирование от индивидуальности; оценочный характер полученных результатов; корректное использование специфических статистических методов сбора, обобщения и анализа данных.

Тема 2. Измерения в социальных исследованиях

Понятие переменной. Виды измерительных шкал: шкала наименований (номинальная шкала), шкала порядка, количественные шкалы интервалов и отношений. Свойства шкал: отношения тождества, порядка, разности, отношения. Дискретные и непрерывные переменные. Представление различных видов шкал в инструментарии исследования. Закрытые и открытые вопросы. Кодирование открытых вопросов. Вторичные измерения: переменные-индикаторы и переменные-индексы. Основные методы построения индексов. Относительные и абсолютные показатели. Реактивность измерения в социальном исследовании. Нерективные методики измерения.

Тема 3. Представление данных социальных исследований

Место статистики в социальном эмпирическом исследовании. Понятие информации. Данные социального исследования как формализованная и структурированная информация об объекте исследования. Этапы формализации информации: определение генеральной совокупности (объекта исследования); построение выборки (выбор единиц наблюдения); операционализация понятий (выбор измеряемых показателей и определение способов их измерения); измерение показателей на единицах наблюдения. Критерии структурирования данных: объект / переменная / время измерения. Структурирование данных в одномоментном исследовании: матрица данных «объект – переменная». Подготовка данных к вводу в компьютер. Кодирование открытых вопросов. Проблема пропущенных значений: легитимные и нелегитимные пропущенные значения. Кодирование пропущенных значений.

РАЗДЕЛ 2. Дескриптивный анализ данных

Тема 4. Одномерные частотные распределения для дискретных и непрерывных переменных

Дескриптивная статистика как средство описания выборочной совокупности. Абсолютная и относительная частота. Одномерное частотное распределение дискретной переменной (номинальной, порядковой, количественной). Одномерное частотное распределение непрерывной переменной: понятие группировки. Виды группировок: типологическая, аналитическая, процентильная. Проблема точных границ интервалов. Типологическая группировка: применение, особенности, правила построения. Аналитическая группировка: применение, особенности, правила построения. Понятие квантиля. Процентильная (квантильная) группировка: применение, особенности, правила построения. Накопленная (кумулятивная) частота: восходящая, нисходящая. Распределение накопленных частот для дискретных и непрерывных переменных. Использование накопленных частот при построении процентильных группировок.

Тема 5. Графическое представление одномерных распределений

Графики как способ визуализации одномерных распределений. Общие требования к графикам. Использование площади фигур в качестве наиболее общего способа представления частот. Шкалы переменных и частот: масштаб, прерывание шкал. Диаграммы для дискретных переменных (построение, особенности использования): круговая диаграмма; диаграммы полос и столбцов; ленточная диаграмма; пиктограмма. Гистограмма и полигон распределения для количественных переменных. Плотность распределения. Правила построения гистограмм и полигонов. Графики накопленных частот (кумуляты): правила построения для дискретных и непрерывных переменных. График интерквартильного диапазона: построение, применение. Статистические карты. Графическое представление временных рядов.

Тема 6. Показатели центра распределения и разброса данных

Понятие центра распределения как разновидности «нормы». «Среднее» как «типичное»: мода (вероятностное среднее). Определение моды для дискретных и сгруппированных переменных. Медиана (ранговое среднее). Определение медианы по распределению накопленных частот. Вычисление медианы для сгруппированных данных. Среднее арифметическое. Вычисление среднего арифметического для сгруппированных данных и оценочных (квазиколичественных) шкал. Среднее арифметическое для дихотомических шкал. Вариационный размах. Среднее квадратическое (стандартное) отклонение и дисперсия, их преимущества перед другими мерами разброса данных. Вычисление стандартного отклонения и дисперсии для сгруппированных данных и оценочных шкал. стандартное отклонение и дисперсия для дихотомических шкал Коэффициент вариации; его использование в сравнительном анализе.

Тема 7. Анализ формы распределения. Стандартизация переменных

Форма эмпирического распределения, ее основные характеристики: модальность, протяженность, симметричность. Классификация распределений в зависимости от количества и расположения мод. Одномодальные распределения: «колокол», J-образное распределение. Бимодальные распределения: «двойной колокол», U-образное распределение. Полимодальное распределение. Анализ природы би- и полимодальности: неоднородность генеральной совокупности. Симметричность одномодального распределения. Правая и левая асимметрия: коэффициент асимметрии. Форма распределения «колокол», коэффициент эксцесса. Анализ формы распределения: использование мер центральной тенденции, показателей разброса данных, графиков. Стандартизация количественных переменных: z -оценка как безразмерная стандартизированная переменная. Распределение и свойства z -оценок.

РАЗДЕЛ 3. Основы статистического вывода

Тема 8. Теория вероятностей как основа статистического вывода

Задачи статистического вывода: оценка неизвестных параметров генеральной совокупности, статистическая проверка гипотез о свойствах генеральной совокупности по данным выборочного исследования. Случайность отбора как предпосылка статистического вывода. Простая случайная выборка. Случайная величина, ее значения: дискретные и непрерывные случайные величины. Выборочное пространство. Случайное событие. Вероятность случайного события. Свойство аддитивности вероятности. Распределение случайной величины. Закон распределения дискретной случайной величины. Функция плотности распределения непрерывной случайной величины. Функция распределения дискретной и непрерывной случайной величины. Свойства закона распределения, функции плотности распределения, функции распределения.

Тема 9. Теоретические распределения, статистические таблицы

Понятие теоретического распределения. Теоретические распределения, наиболее часто используемые в анализе данных социальных исследований: Гаусса (нормальное), Стьюдента, Фишера, хи-квадрат. Нормальное распределение Гаусса $N(\mu, \sigma)$, его параметры, свойства. Стандартное нормальное распределение $Z(0,1)$. Таблица стандартного нормального распределения. Использование таблицы стандартного нормального распределения для работы с произвольными нормальными распределениями. Распределение Стьюдента $t(df)$, его параметр, свойства, таблица распределения. Распределение хи-квадрат $\chi^2(df)$, его параметр, свойства, таблица распределения. Распределение Фишера $F(df_1, df_2)$, его параметры, свойства, таблицы распределения.

Тема 10. Оценивание параметров генеральной совокупности.
Репрезентативность выборки

Параметры генеральной совокупности и выборочные статистики. Ошибка выборки. Случайная и систематическая составляющие ошибки выборки, их источники. Выборочное распределение статистики. Следствие из Центральной предельной теоремы. Точечные оценки параметров генеральной совокупности. Свойства точечных оценок: несмещенность, эффективность, состоятельность. Точечное оценивание методом моментов. Интервальное оценивание параметров генеральной совокупности. Стандартная ошибка выборки. Доверительная вероятность, ее стандартные значения: 0,99; 0,95; 0,9. Доверительный интервал. Построение доверительных интервалов для математического ожидания, доли единицы, дисперсии. Доверительный интервал для случайной ошибки выборки. Показатели репрезентативности выборки: предельно допустимая ошибка и доверительная вероятность. Минимальный объем простой случайной репрезентативной выборки. Поправка на конечный объем генеральной совокупности. Конкретный и субъективный характер репрезентативности. Представление данных о репрезентативности выборки в научных работах и прессе.

Тема 11. Статистическая проверка гипотез

Понятие и структура статистической гипотезы. Назначение нулевой гипотезы H_0 и альтернативных гипотез H_1 . Односторонние и двусторонние альтернативные гипотезы. Верные и ошибочные решения при проверке статистической гипотез. Ошибка первого рода, ее вероятность, уровень значимости. Ошибка второго рода, ее вероятность. Критерий нулевой гипотезы, мощность критерия. Статистика критерия, ее распределение, свойства. Процедура проверки статистической гипотезы. Критическое значение и критическая область. Односторонние и двусторонние критические области. Использование статистических таблиц при проверке гипотез. Проверка статистических гипотез с помощью компьютера: p -значение. Проверка гипотез о равенстве математических ожиданий (t -критерий); вероятностей положительного ответа (Z -критерий); дисперсий (F -критерий, критерий хи-квадрат); распределений вероятностей (критерий хи-квадрат).

РАЗДЕЛ 4. Анализ статистических связей

Тема 12. Понятия статистической связи и независимости

Понятия статистической связи и статистической независимости. Природа статистической связи. Парная и множественная связь между переменными. Ненаправленная связь (для номинальных переменных). Прямая и обратная связь (для порядковых и количественных переменных). Понятие меры связи; требования к мерам связи. Корреляционная и причинная связь. Определение причинной связи. Формальные критерии причинности (каузальности). Зависимые и независимые переменные.

Тема 13. Частотные модели сопряженности

статистической связи. Таблица

Таблица сопряженности как инструмент представления совместного распределения двух переменных. Элементы таблицы сопряженности, правила ее заполнения. Маргинальные частоты. Определение статистической зависимости и независимости по таблице сопряженности. Частотная модель связи: критерий хи-квадрат статистической независимости строк и столбцов таблицы. Эмпирические и теоретические частоты. Проверка гипотезы о наличии связи между строками и столбцами таблицы сопряженности. Измерение силы (тесноты) связи. Меры связи, основанные на критерии хи-квадрат: коэффициенты контингенции (ϕ), Чупрова (T), Крамера (V), преимущества и недостатки коэффициентов, их интерпретация. Таблицы сопряженности размерности 2×2 . Понятия абсолютной (двусторонней) и полной (односторонней) связи; прямой и обратной связи для таблиц 2×2 . Меры связи для таблиц 2×2 : коэффициенты Юла Q и контингенции Φ . Использование коэффициентов Q и Φ при анализе и интерпретации связи двух дихотомических переменных. Теоретико-информационные меры связи для номинальных таблиц: λ . Меры связи для упорядоченных таблиц сопряженности: γ Гудмана и Краскалла, d Сомера, τ_b Кендалла. Анализ стандартизированных остатков

Тема 14. Меры связи, основанные на рангах

Понятие ранговой корреляции, особенности ее применения в прикладных задачах. Прямая и обратная ранговая связь. Коэффициенты ранговой корреляции r_s Спирмана и τ Кендалла, интерпретация их значений. Понятие связанных рангов. Коэффициенты Спирмана и Кендалла для связанных рангов. Коэффициенты Спирмана и Кендалла для упорядоченных таблиц сопряженности. Прямая и обратная связь в упорядоченных таблицах. Интерпретация коэффициентов ранговой корреляции для таблиц сопряженности. Коэффициент множественной ранговой корреляции, его вычисление и интерпретация.

Тема 15. Линейная статистическая модель парной связи.

Корреляционный и регрессионный анализ

Линейная статистическая модель парной связи. Диаграмма рассеяния как способ графического представления совместного распределения двух количественных переменных. Анализ линейности, направления и тесноты связи двух переменных по диаграмме рассеяния. Ковариация как мера совместного рассеяния двух переменных, ее свойства. Коэффициент линейной корреляции Пирсона, его свойства. Связь коэффициента Пирсона с коэффициентом Φ для дихотомических переменных и ранговым коэффициентом Спирмана. Проверка гипотезы о статистической значимости коэффициента линейной корреляции. Регрессионный анализ как метод исследования причинной линейной связи. Зависимая и независимая переменные в уравнении линейной регрессии. Построение уравнения парной линейной регрессии методом наименьших квадратов. Вычисление и

интерпретация параметров уравнения регрессии. Анализ остатков регрессии. Модель с разделением дисперсии: дисперсия общая, объясненная, остаточная. Коэффициент детерминации, его интерпретация в качестве доли объясненной дисперсии.

Тема 16. Нелинейные модели парной связи

Нелинейные модели парной связи, их вид на диаграмме рассеяния. Разделение дисперсии зависимой переменной: общая, внутригрупповая и межгрупповая (объясненная) дисперсия. Корреляционное отношение как показатель тесноты нелинейной связи между зависимой и независимой переменными. Преимущества и недостатки корреляционного отношения. Нелинейные модели с использованием логарифмических и экспоненциальных функций.

Тема 17. Основы планирования эксперимента. Дисперсионный анализ

Метод эксперимента в социальных исследованиях. Зависимая и независимые переменные (факторы); требования к ним. Прямые эффекты факторов и эффекты взаимодействия. Классический трехстадийный эксперимент: формирование групп; проведение эксперимента; оценивание результатов. Характеристики эксперимента: число факторов, характер факторов (выделяемые, контролируемые), число групп, схема (внутригрупповая, межгрупповая). Виды экспериментов: лабораторный, естественный (квазиэксперимент), полевой. Представление результатов экспериментального исследования: план эксперимента. Модель дисперсионного анализа: общая, объясняемая (межгрупповая) и остаточная (внутригрупповая) дисперсия. Разложение общей суммы квадратов на меж- и внутригрупповую суммы квадратов. Числа степеней свободы для сумм квадратов. Однофакторный дисперсионный анализ; его гипотеза. Проверка нулевой гипотезы об отсутствии различий между средними значениями зависимой переменной в группах, образованных различными градациями переменной-фактора: t -критерий; F -критерий. Двухфакторный дисперсионный анализ. Главные эффекты факторов и эффект взаимодействия как источники межгрупповой дисперсии. Гипотезы двухфакторного дисперсионного анализа. Проверка гипотез: F -критерий.

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

ЛИТЕРАТУРА

ОСНОВНАЯ

1. *Елисеева И. И., Юзбашев М. М.*. Общая теория статистики. – М.: Финансы и статистика, 2005. – 657 с.
2. *Крыштановский А.О.* Анализ социологических данных. – М.: ГУ ВШЭ, 2006. – 281 с.
3. *Наследов, А. Д.* Математические методы психологического исследования. Анализ и интерпретация данных: учебник. – СПб: Речь, 2004г. – 1-е изд., 2006г. – 2-е изд., 2008г. – 3-е изд., 2011г. – 4-е изд. – 388 с.
4. *Наследов, А. Д.* IBM SPSS Statistics 20 и AMOS: Профессиональный статистический анализ данных». СПб: Питер, 2013г.
5. *Толстова Ю. Н.* Математические модели в социологии. М.: ГУ ВШЭ, 2007, – 244 с.

ДОПОЛНИТЕЛЬНАЯ

1. *Афифи А., Эйзен С.* Статистический анализ: подход с использованием ЭВМ. – М.: Финансы и статистика, 1982.
1. *Бююль А., Цёфель П.* SPSS: искусство обработки информации. СПб: Питер, 2002.
2. *Гласс Дж., Стенли Дж.* Статистические методы в педагогике и психологии. М.: Прогресс, 1976. – 495 с.
3. *Вихалемм П.* Эксперимент в социологическом исследовании. // Методы сбора информации в социологическом исследовании. Кн.2. М., 1990.
4. *Закс Л.* Статистическое оценивание. М., 1976.
5. *Кендалл М. Дж., Стьюарт А.* Статистические выводы и связи. М., 1973.
6. *Малхотра Н.* Маркетинговые исследования: практическое руководство. М., 2007.
7. *Паниотто В. И., Максименко В. С.* Количественные методы в социологических исследованиях. Киев: Наукова думка, 1982. – 272 с.
8. *Пацюрковский В. В., Петрова А. И., Пацюрковская В. В.* Использование SPSS в социологии. Ч. 1. Ввод и контроль данных. М., 1998.
9. *Сатаров Г. А.* Математика в социологии: стереотипы, предрассудки, заблуждения. // Социологические исследования, 1986. N3.
10. Статистические методы анализа информации в социологических исследованиях. М., 1979.
11. *Терещенко О. В.* Прикладная статистика для социальных наук: Компьютерный практикум для студентов гуманитарных специальностей. Минск: БГУ, 2002, – 93с.
12. *Толстова Ю. Н.* Измерение в социологии. М.: Инфра-М, 1998, – 221 с.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ И ВЫПОЛНЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Самостоятельная работа студентов организуется в соответствии с учебным планом специальности на основании действующего в вузе Положения о самостоятельной работе студентов.

Самостоятельную работу студентов по дисциплине «Прикладная статистика» в аудитории рекомендуется осуществлять в форме решения задач, выполнения контрольных работ и заданий по анализу данных в компьютерном классе. Внеаудиторная самостоятельная работа студентов предполагает выполнение домашних заданий по решению задач, подготовку к практическим занятиям и экзамену.

ПЕРЕЧЕНЬ РЕКОМЕНДУЕМЫХ СРЕДСТВ ДИАГНОСТИКИ РЕЗУЛЬТАТОВ УЧЕБНОЙ ДЕЯТЕЛЬНОСТИ

Для диагностики результатов учебной деятельности рекомендуется использовать оценки домашних практических заданий, контрольных работ, письменных отчетов по выполнению заданий в компьютерном классе, тестирование. Рекомендуемая форма отчетности – экзамен, который может быть проведен как в устной, так и в письменной форме.