Приложение 10 ОП ВО

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Статистика в сфере публичной политики (Statistics for Public Policy)
наименование дисциплины

Автор: доцент кафедры международного менеджмента, PhD Нейштадт И. В.
Код и наименование направления подготовки, профиля: 41.04.04 Политология
Квалификация (степень) выпускника: Магистратура
Форма обучения: Очная

Цель освоения дисциплины:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК-3	Углубленное знание общих и специальных методов современной политической науки, уверенное владение навыками применения методологии политической науки к анализу современных политических процессов	ПК-3.1	Знание общих и специальных методов политического анализа.

План курса:

№	Наименование	Содержание
Тема 1	Introduction to Statistics & Research Design	• How can we ask and answer meaningful empirical questions? • Randomized Experiments and Causality, Observational Studies • Sampling Design, Surveys and Questionnaires ТА: Introduction to STATA (Log files, Data Editor, Recoding Variables)
Тема 2	Exploratory Data Analysis—Single Variable	• Classification of Variables, Frequency Distributions, Graphical Displays (Bar Charts, Pie Charts, Histograms) • Measures of Central Tendency (Mean, Median, Mode), Measures of Dispersion (Quartiles, Boxplots, Variance, Standard Deviation) • Density curves, Standard Normal Distribution (Z Table, Normal Quantile Plots) STATA: Frequency tables, Charts and Graphs, Summary Statistics
Тема 3	Exploratory Data Analysis—Relationships between variables	• Association between categorical explanatory variable and quantitative response variable • Association between 2 categorical variables • Association between 2 quantitative variables: Scatterplots, Correlation STATA: Means tables, Crosstabs, Scatterplots
Тема 4	Exploratory Data Analysis—Simple Regression	• Simple Linear Regression (Least Squares Criterion, Calculating and Interpreting regression coefficients) • Predicted Values and Residuals, R-square, Outliers, Lurking Variables, Causation STATA: Simple Regression, residual analysis

№	Наименование	Содержание
Тема 5	Concepts and Applications in Probability	• Basic Definitions, Compound Events, Joint Probabilities, Conditional Probabilities • Statistical Independence, Contingency Probability Tables • Random Variables (Discrete versus Continuous), Mean and Variance of Random Variables, Probability Distributions STATA: Contingency probability tables (crosstabs)
Тема 6	Sampling Distributions I	• Sampling Distribution for Counts (Calculating Binomial Probabilities), Sampling Distribution for proportions, Normal approximation • Sampling Distribution of the Mean, Central Limit Theorem.
Тема 7	Basics of Statistical Inference	• Student's t-distribution, Confidence Interval around a single Mean • Hypothesis testing (Null and Alternative, One versus two-sided tests, p-values) STATA: One-Sample t-test
Тема 8	Inference for two sample means, a single proportion, and two proportions	• Inference for Difference in Means (Independent samples, Pooled Variance), Confidence Interval around a Difference in Means, Hypothesis testing • Inference for Single Proportion (Confidence Intervals, Hypothesis testing) • Inference for Difference between two proportions (CI's, Hypothesis testing) STATA: Independent Samples t-test, Test for a single proportion, Two-sample test of proportion
Тема 9	Contingency Tables & Inference for Simple Linear Regression	• Contingency Tables Reviewed, Chi-Square test of statistical independence • Simple Linear Regression, Goodness of fit, Gauss-Markov Assumptions • Hypothesis testing of slope coefficient (Standard error, Confidence Intervals, p-values) STATA: Crosstabs, Chi-Square, Simple regression estimates, assess statistical significance, calculate predicted values and residuals, goodness of fit
Тема 10	Multiple Linear Regression I	• Review Randomized Experiments versus Observational Studies, Confounding Factors • Fitting the model, Check of regression assumptions, Model fit (F-test, R-square) STATA: Multiple regression estimates, use residuals to check regression assumptions
Тема 11	Multiple Linear Regression II	• Dummy Independent Variables (Same slopes, different intercepts) • Interaction terms (Different slopes, different intercepts) • Categorical Independent Variables STATA: Creating and using dummy independent variables, creating and using interaction terms

Формы текущего контроля и промежуточной аттестации:

Примеры задач в классе:

Examples of class problems:

Problem 1. (10 pts.) Poker hands.

After one-pair, the next most common hands are two-pair and three-of-a-kind:

Two-pair: Two cards have one rank, two cards have another rank, and the remaining card has a third rank. e.g. {2♥, 2♠, 5♥, 5♣, K♦}

Three-of-a-kind: Three cards have one rank and the remaining two cards have two other ranks. e.g. {2♥, 2♠, 2♣, 5♣, K♦}

Calculate the probability of each type of hand. Which is more likely?

Problem 2. (10 pts.) Non-transitive dice.

In class we worked with non-transitive dice:

Red: 3 3 3 3 3 6; Green: 1 4 4 4 4 4; White: 2 2 2 5 5 5.
Finish making your non-transitive dice.
(a) In class you found the probability that red beats white. Find the probability that white beats green and the probability that green beats red.
Can you line the dice up in order from best to worst? (Hint: this is why these are called 'non-transitive'.)
(b) Suppose you roll two white dice against two red dice. What is the probability that the sum of the white dice is greater than the sum of the red dice?

Problem 3. (20 pts.) Birthdays: counting and simulation.
Ignoring leap days, the days of the year can be numbered 1 to 365. Assume that birthdays are equally likely to fall on any day of the year. Consider a group of n people, of which you are not a member. An element of the sample space Ω will be a sequence of n birthdays (one for each person).
(a) Define the probability function P for Ω..
(b) Consider the following events:
A: "someone in the group shares your birthday"
B: "some two people in the group share a birthday"
C: "some three people in the group share a birthday"
Carefully describe the subset of Ω that corresponds to each event.
(c) Find an exact formula for P(A). What is the smallest n such that P(A) > .5?
(d) Justify why n is greater than 365 without doing any computation. (We are looking for 2 a short answer giving a heuristic sense of why this is so.)
(e) Use R simulation to estimate the smallest n for which P(B) > .9. For these simulations, let the number of trials be 10000.
For this value of n, repeat the simulation a few times to verify that it always gives similar results. Using 10000 trials you saw very little variation in the estimate of P(B). Try this again using 30 trials and verify that the estimated probabilities are much more variable.
(f) Find an exact formula for P(B).
(g) Use R simulation to estimate the smallest n for which P(C) > .5. Again use 10000 trials. You will find that two values of n are equally plausible values. You may pick either one for your answer.

Типовые оценочные средства

Письменный экзамен
Written exam.
1. There are 3 arrangements of the word DAD, namely DAD, ADD, and DDA. How many arrangements are there of the word PROBABILITY?
2. There are six men and seven women in a ballroom dancing class. If four men and four women are chosen and paired off, how many pairings are possible?
3. Let A and B be two events. Suppose the probability that neither A or B occurs is 2/3. What is the probability that one or both occur?
4. Let C and D be two events with P(C) = 0.25, P(D) = 0.45, and P(C ∩ D) = 0.1. What is P(Cc ∩ D)?
5. Suppose you are taking a multiple-choice test with c choices for each question. In answering a question on this test, the probability that you know the answer is p. If you don't know the answer, you choose one at random. What is the probability that you knew the answer to a question, given that you answered it correctly?
6. Two dice are rolled.

A = ‘sum of two dice equals 3’

B = ‘sum of two dice equals 7’

C = ‘at least one of the dice shows a 1’

(a) What is P(A|C)?
(b) What is P(B|C)?
(c) Are A and C independent? What about B and C?

7. A multiple choice exam has 4 choices for each question. A student has studied enough so that the probability they will know the answer to a question is 0.5, the probability that they will be able to eliminate one choice is 0.25, otherwise all 4 choices seem equally plausible. If they know the answer they will get the question right. If not they have to guess from the 3 or 4 choices.

As the teacher you want the test to measure what the student knows. If the student answers a question correctly what’s the probability they knew the answer?

8. Suppose that P(A) = 0.4, P(B) = 0.3 and P((A ∪ B)C) = 0.42. Are A and B independent?

9. Suppose that X takes values between 0 and 1 and has probability density function 2x. Compute Var(X) and Var(X2).

10. Compute the expectation and variance of a Bernoulli(p) random variable.

Основная литература:

D. Moore, G. McCabe, and B. Craig. Introduction to the Practice of Statistics.7th Ed (2012)