NEW YORK CITY COLLEGE OF TECHNOLOGY
The City University of New York

DEPARTMENT:	Mathematics
COURSE:	MAT 1372
TITLE:	Statistics with Probability
DESCRIPTION:	Topics covered include sample spaces and probabilities, discrete probability distributions (Binomial, Hypergeometric, Poisson), expectation and variance, continuous probability distributions (Normal, Student, Chi-Square), confidence intervals, hypothesis testing, correlation and regression. Spreadsheets are used throughout the semester.
TEXTS:	OpenIntro Statistics (4th edition, 2019)
CREDITS:	3 (2 class hours, 2 lab hours)
PRE- or COREQUISITES:	MAT 1375
	Prepared by Prof. Ezra Halleck (Spring 2021)

A. Testing/Assessment Guidelines:
The following exams should be scheduled:
1. A one session exam at the end of the First Third.
2. A one session exam at the end of the Second Third.
3. A one session Final Examination.

Multiple short projects or one long project should be part of the course.

B. Statistical software will be used throughout the semester. Students should plan to spend time with tutorials, especially those unfamiliar with the software chosen by the instructor.

Learning Outcomes: Upon completion of the course, students will be able to

Course Learning Outcome	General Education Learning Outcome	Flexible Core Scientific World Learning Outcome
Collect, organize, group and display data in a table and in graphs, analyze for shape and skewness and interpret within context.	IL(IC): present data in appropriate ways to best illustrate a point. W(Context): identify a context of a situation in order to choose an appropriate representation of data.	Gather, interpret, and assess information from a variety of sources and points of view.
Compute and interpret measures of central location, variability and position, such as mean, median, mode, standard deviation and percentiles. In communication, make effective use of both probability and statistics terminology.	IL(CE): connect relevant experience and academic knowledge. R(Context): interpret statistics in the context of a problem. W(L): express solutions using appropriate terminology.	Identify and apply the fundamental concepts and methods of a discipline or interdisciplinary field exploring the scientific world, including, but not limited to computer science, history of science, life and physical sciences, linguistics, logic, mathematics, psychology, statistics, and technology-related studies.
Understand how inference relies on a probabilistic analysis of sampling. Apply the Central Limit Theorem to find the mean and standard deviation of a sampling distribution as well as its shape.	IL(R): reflect as a learner building on prior experience to develop or understand experiments carried out in this context. R(A): describe an experiment that would test a given probability. W(CP): explain the context and purpose of an experiment in words or orally.	Demonstrate how tools of science, mathematics, technology, or formal analysis can be used to analyze problems and develop solutions.
Solve a problem by setting up a hypothesis test and using either the rejection region or p-value approach.	I(T): connect ideas of probability and statistics to test hypotheses and estimate level of significance of such a test. R(I): interpret the results of a hypothesis test in terms of the original problem. W(AS): analyze a real situation and form, carry out and interpret the results of a hypothesis test. W(F,C): produce a written report of a hypothesis test with appropriate labeling and logical flow of supportive evidence to support conclusions.	Evaluate evidence and arguments cri tically or analytically. Produce well-reasoned written or oral arguments using evidence to support conclusions.
Provide real-world examples modeled by binomial, hypergeometric, Poisson, normal, student-t and chi-squared distributions.	IL(CD): apply statistical problems in real life problems. R(Context): find the appropriate model for a real-life situation.	Produce well-reasoned written or oral arguments using evidence to support conclusions.
Simulate a probability distribution as well as the process of going from a population to a sample.	IL(T): adapt ideas of real experiments to the context of computer experiments. W(SE): perform computer experiments to support ideas and arguments.	Articulate and evaluate the empirical evidence supporting a scientific or formal theory.

New York City College of Technology Policy on Academic Integrity

Students and all others who work with information, ideas, texts, images, music, inventions, and other intellectual property owe their audience and sources accuracy and honesty in using, crediting, and citing sources. As a community of intellectual and professional workers, the College recognizes its responsibility for providing instruction in information literacy and academic integrity, offering models of good practice, and responding vigilantly and appropriately to infractions of academic integrity. Accordingly, academic dishonesty is prohibited in The City University of New York and at New York City College of Technology and is punishable by penalties, including failing grades, suspension, and expulsion. The complete text of the College policy on Academic Integrity may be found in the catalog.

session		Section	p.	Exercises	concepts/skill
1		Introduction to R (or Excel)			
	1 Introduction to data 7				
2		*1.1 Case study: using stents to prevent strokes	9	1.1	contingency table, control, placebo
2		1.2 Data basics	12	1.3,5,6,7	variable types, explanatory, response
2		1.3 Sampling principles and strategies	22	1.13,14,16	sample method, population, inference
2		1.4 Experiments	32	1.29,30	observation vs experiment
	2 Summarizing data 39				
3		2.1 Examining numerical data	41	2.1-10	linear, nonlinear, no relation, sample mean, grouped data, histogram, box plot, mean, median, skew, IQR,
4		*2.2 Considering categorical data	61	2.21-23	bar, pie, contingency table, mosaic
4		*2.3 Case study: malaria vaccine	71	2.26	simulation, small sample inference
	3 Probability 79				
5		3.1 Defining probability	81	3.3-7odd,8,10	disjoint, Venn diagram, independence, complement, Law of large #s
6		3.2 Conditional probability	95	3.13,16,18,19,21	joint/conditional probabilities, trees, independence, Bayes Thm
7		3.4 Random variables	115	3.31-34,36	expectation, std dev
8		3.5 Continuous distributions	125	3.37	proportions of ranges from histogram
8		opentext.bc: continuous distributions		prob. end of section	discrete vs continuous
8		openstax: continuous distributions		72, 73	
8		opentext.bc: uniform distribution		rand#/wtLoss/sbwy	uniform dist, area under curve
8		openstax: uniform distribution		75-85 odd	
	4 Distributions of random variables 131				
9		4.1 Normal distribution	133	4.1-9 odd, 10	std, non-std, percentile, inverse
10		4.3 Binomial distribution	149	4.17,19,21,24,26	
11		*3.3 Sampling from a small population	112	3.24,26,28	counting, combinations, complement
11		*opentext.bc: hypergeometric distribution			
11		*openstax: hypergeometric distribution		114,116	probabilities, mean, std dev.
12		4.5 Poisson distribution	163	4.31-34	justify using or not Poisson, calc probs

session		Section	p.	Exercises	concepts/skill
	5 Foundations for inference 168				
13		5.1 Point estimates and sampling variability	170	5.1-5	identify parameter (mean/proportion), p vs phat and standard error
14		5.2 Confidence intervals for a proportion	181	5.7-11	margin of error
15		5.3 Hypothesis testing for a proportion	189	5.15-23 odd	set up hypothesis, test claim, wk bkwds
	6 Inference for categorical data 206				
16		6.1 Inference for a single proportion	208	6.10,11,12,14,15	given margin of error, find sample size
17		*6.2 Difference of two proportions	217	6.17,18,19,27,28	find conf interval, test hyp diff=0
18		6.3 Testing for goodness of fit using chi-square	229	6.31-34	chi-sq properties and GOF hyp testing.
19		6.4 Testing for independence in two-way tables	240	6.35-6.38	expected values and hyp test
	7 Inference for numerical data 249				
20		7.1 One-sample means with the t-distribution	251	7.2,3,5,7,8	graph,hyp test, wk bkwds from interval
21		*7.2 Paired data	262	7.15,16,17,19,20	determine if paired, hyp test
22		*7.3 Difference of two means	267	7.24-32 even	hyp test or conf interval, criterion
	8 Introduction to linear regression 303				
23		8.1 Fitting a line, residuals, and correlation	305	8.1-9 odd,13,14	vis. residuals, calc/match correlation
24		8.2 Least squares regression	317	8.17,19,21,24,25	lin reg good model?, find/interp slope
25		*8.3 Types of outliers in linear regression	328	8.27,29,30	identify/classify outliers, leverage
25		*8.4 Inference for linear regression	331	8.31,34	hyp test, interpret coef of det
3 sessions for midterm exams or project presentations					
1 session for review					
1 session for final exam					
* optional					