School of Mathematics and Statistics

Te Kura Mātai Tatauranga

STAT 292 Assignment 1: Due Thursday, 16 March 2023 at 11:59 PM

Note: Your assignment can be typed or handwritten (and scanned). Be sure to submit your assignment as a PDF and follow the instructions specified on the course Nuku page. Where calculations are performed in R, you must include relevant code and output with your answer to receive credit.

Assignments that are submitted late will receive a mark of 0 unless illness, bereavement or other substantial causes occur and have been discussed with the course coordinator.

1. (3 marks)

   Which of the following variables are categorical, and why?

   a. Inland Revenue Department (IRD) tax number.
   b. Level of agreement with the federal government's COVID response (1 = "Strongly disagree", 2 = "Disagree", 3 = "Neither agree nor disagree", 4 = "Agree", 5 = "Strongly agree").
   c. Grade point average (on a scale of 0.0 to 4.0).
   d. Speed limit (in kilometers per hour).
   e. COVID incubation time (in days).
   f. Phone number.

2. (15 marks)

   Data reported by Stats NZ in 2022 estimated that, in the financial year ending in June 2021, 16% of children in New Zealand lived in poverty (as defined by the household in which they lived having after-housing-costs [AHC] equivalised disposable income of less than half the median AHC income for 2021). Consider a random sample of 60 children who lived in New Zealand in the financial year ending in June 2021, and suppose that the number of these children living in poverty in the financial year ending in June 2021 can be represented by a random variable following a binomial distribution.

   a. Cleary explain what we are assuming about these 60 children in representing the number of these children living in poverty in the financial year ending in June 2021 by a binomial distribution. Provide an example of when this assumption would likely be violated. (Your answer must clearly refer to the situation described in the problem.) (3 marks)
   b. What is the mean number of these 60 children that would be expected to have lived in poverty in the financial year ending in June 2021? What are the corresponding variance and standard deviation? (3 marks)

c. Using `R`, calculate the probability (to at least 3dp) that exactly 15 of these 60 children lived in poverty in the financial year ending in June 2021. (Be sure to show your `R` code and output.) (2 marks)

d. What is the probability (to at least 4dp) that more than 10 of these 60 children lived in poverty in the financial year ending in June 2021? Calculate this probability

   i. exactly using `R` and the binomial distribution (be sure to include your `R` code and output) and

   ii. by hand using a normal approximation and the standard normal probability table.

   (7 marks)

3. (4 marks)

Suppose the New Zealand government wants to estimate the proportion of children who live in poverty for the financial year ending in June 2022. They would like to produce a 90% confidence interval for this proportion. Previous studies have found that the proportion of children living in poverty in New Zealand is almost certainly no more than 30%. Using this information, find the most conservative minimum sample size required to produce a confidence interval with an approximate margin of error of 0.02.

4. (13 marks)

Consider the following data showing random samples of children from two different regions of New Zealand (Southland, Bay of Plenty) and the numbers of these children who lived in poverty in the financial year ending in June 2021. Relevant data are shown in the table below.

Region	Number of children who did not live in poverty	Number of children who lived in poverty
Southland	500	55
Bay of Plenty	500	95

a. Produce both a standard and an Agresti-Coull 95% confidence interval (to at least 3dp) for the proportion of children from Southland who lived in poverty in the financial year ending in June 2021. (5 marks)

b. Test whether the proportions of children who lived in poverty in the financial year ending in June 2021 are different for the Southland and Bay of Plenty regions using a test for proportions. Be sure to report the hypotheses, test statistic, p-value, and your conclusion at the $\alpha = 0.05$ significance level. Your working should be done by hand, and the p-value should be calculated using the standard normal probability table. (8 marks)

5. (15 marks)

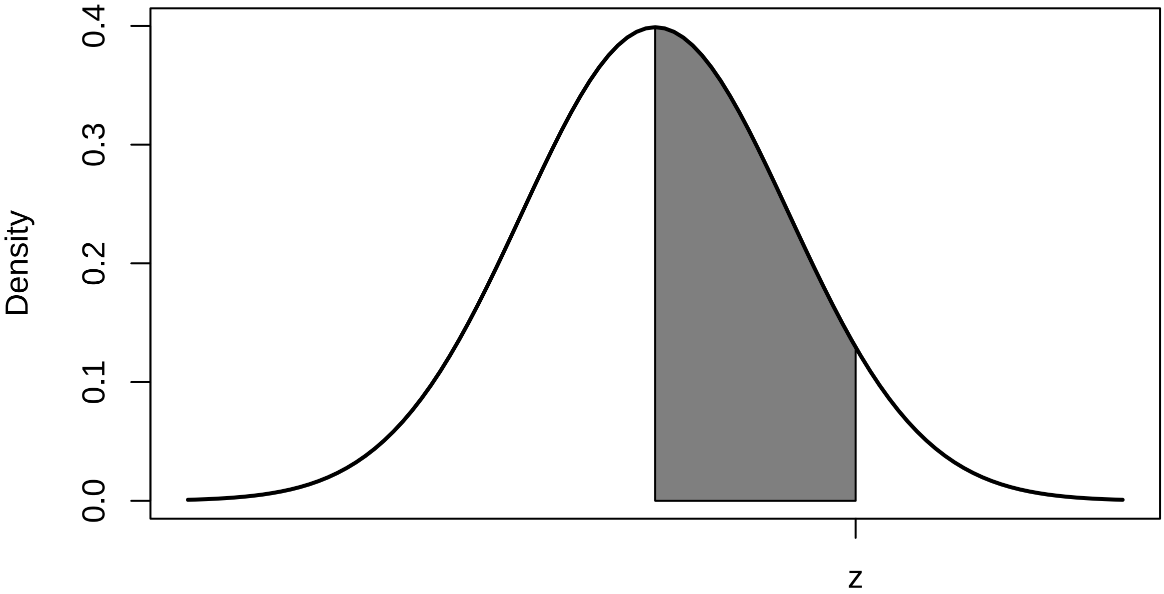
The following data shows the number of flight cancellations over the first 69 days in 2015 for Gerald R. Ford International Airport.

Number of cancellations (r)	Observed frequency (f_r)
0	29
1	17
2	9
3	4
4	3
5	2
6	0
7	1
8	0
9	4

a. Find the mean number of flight cancellations per day (to at least 3dp).
(3 marks)

b. Test whether the number of flight cancellations per day is consistent with a Poisson distribution. Be sure to clearly state the null and alternative hypotheses, present the test statistic and its distribution under the null hypothesis, and report the p-value and your conclusion at the $\alpha = 0.05$ significance level.
(12 marks)

Standard Normal Probability Table



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
—	—	—	—	—	—	—	—	—	—	—
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
—	—	—	—	—	—	—	—	—	—	—
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
—	—	—	—	—	—	—	—	—	—	—
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000