TELFER

VOTRE LIEN AVEC CE QUI COMPTE — CONNECTS YOU TO WHAT MATTERS

ADM 2303, Fall 2016

STATISTICS FOR MANAGEMENT I

FINAL EXAM DECEMBER

GIVEN NAME .... FAMILY NAME .... Student #.................... Section ....

1. Books and notes **are not** permitted. **One** sheet of notes, 8.5”x11”, is allowed as per instructions in course outline.
2. Calculators **are** permitted.
3. Use the space on the question paper for rough work.

STATISTICAL TABLES ARE PRINTED AT THE END OF THIS EXAM

USE THE QUESTION PAPER FOR ROUGH WORK

SHADE THE APPROPRIATE BOX ON THE CODING SHEET

THE QUESTION PAPER CAN BE USED FOR CALCULATIONS.

Statement of Academic Integrity

The School of Management does not condone academic fraud, an act by a student that may result in a false academic evaluation of that student or of another student. Without limiting the generality of this definition, academic fraud occurs when a student commits any of the following offences: plagiarism or cheating of any kind, use of books, notes, mathematical tables, dictionaries or other study aid unless an explicit written note to the contrary appears on the exam, to have in his/her possession cameras, radios (radios with head sets), tape recorders, pagers, cell phones, or any other communication device which has not been previously authorized in writing.

Statement to be signed by the student:

I have read the text on academic integrity and I pledge not to have committed or attempted to commit academic fraud in this examination.

Signed:_______________________________________

Note: a paper without that signed statement will not be graded and will receive a final exam grade of zero.

Questions 1-5. Hepatitis C in British Columbia. On World Hepatitis Day, July 28, 2014, the BC Centre for Disease Control stated that 3% of British Columbians (about 130,000 people) were infected with hepatitis C. Although there is no vaccination against hepatitis C, it can be cured with a daily combination pill that cures 95% of patients. We survey a sample of 2500 randomly selected British Columbians living in the city of Burnaby as to whether they have hepatitis C.

1) What is the number of people that we would expect to find in our sample that have hepatitis C?

a) 64
b) 75
c) 128
d) 150
e) It depends on the population of Burnaby

2) What is the standard deviation of the proportion of people in our sample that have hepatitis C?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.1
d) between 0.1 and 0.2
e) between 0.2 and 1.0

3) We find that the proportion of people in our sample with hepatitis C is 2.56%. What is the probability that the proportion would be as low as 2.56% (or lower) if the true proportion for the whole of Burnaby is 3%, i.e. the same as for British Columbia as a whole?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.05
d) between 0.05 and 0.2
e) between 0.2 and 1.0

4) What conditions do we need to check in order to answer question 3?

a) the expected number of people in our sample with hepatitis C is > 10
b) the expected number of people in our sample without hepatitis C is > 10
c) the total population of Burnaby is > 25,000
d) both a) and b)
e) all the above

5) Can you conclude that Burnaby has a lower proportion of people with hepatitis C than the whole of British Columbia?

a) Yes, because 0.0256 is less than 0.03
b) Yes, because the probability in question 3 is > 0.05
c) No, because the probability in question 3 is > 0.05
d) Yes, because the probability in question 3 is < 0.05
e) No, because the probability in question 3 is < 0.05

Questions 6-10. Gender Gap. Each year the World Economic Forum measures the differences between men and women with regard to four factors (i) health and survival, (ii) educational attainment, (iii) economic participation and opportunity and (iv) political empowerment. These four measures are combined into one score, which we will call the gender equality score. A higher score implies more gender equality. The results for the top 20 ranked countries for 2014 are:

Iceland	0.8594	Switzerland	0.7798
Finland	0.8453	Germany	0.778
Norway	0.8374	New Zealand	0.7772
Sweden	0.8165	Netherlands	0.773
Denmark	0.8025	Latvia	0.7691
Nicaragua	0.7894	France	0.7588
Rwanda	0.7854	Burundi	0.7565
Ireland	0.785	South Africa	0.7527
Philippines	0.7814	Canada	0.7464
Belgium	0.7809	United States	0.7463

[The total of the gender equality scores for the 20 countries given above is 15.721]

6) If you were to draw a histogram of these results, how many bins (groups, bars) would you use. You are not asked to actually draw the histogram, just say how many bins (groups, bars) you would use.

a) 4
b) 5
c) 6
d) 10
e) either a) or b)

7) Calculate the mean of the distribution of the gender equality score for the countries shown. Choose the closest option below.

a) 0.78
b) 0.79
c) 0.8
d) 0.81
e) 0.82

8) Calculate the median of the distribution of the gender equality score for the countries shown. Choose the closest option below.

a) 0.78
b) 0.79
c) 0.8
d) 0.81
e) 0.82

9) What do your answers to 7) and 8) say about the shape of the distribution of gender gap for the countries shown?

a) It is exactly symmetric
b) It is skewed to the right
c) It is skewed to the left
d) The distribution has more than one mode
e) We cannot conclude anything about the shape of the distribution

10) Which country corresponds to the 7^{th} percentile of the distribution of gender gap for the countries shown?

a) United States; b) Canada; c) Iceland; d) Finland;
e) The 7^{th} percentile is in between two countries.

Questions 11-16. A market research firms conducts studies regarding the success of new products. The company is not always perfect in predicting the success. Suppose that there is a 65% chance that any new product would be successful (and a 35% chance that it would fail). In the past, for all new products that ultimately were successful, 80% were predicted to be successful (and the other 20% were inaccurately predicted to be failures). Also, for all new products that were ultimately failures, 70% were predicted to be failures (and the other 30% were inaccurately predicted to be successes).

11) What is the probability that a new product would be a success?

a) between 0.65 and 0.70
b) between 0.70 and 0.75
c) between 0.80 and 0.85
d) between 0.75 and 0.80
e) This probability cannot be obtained

12) For a sample of 6 randomly selected successful new products, what is the probability that none was mistakenly predicted to be a failure?

a) < 0.10
b) between 0.10 and 0.20
c) between 0.20 and 0.30
d) between 0.30 and 0.40
e) > 0.40

13) For any randomly selected new product, what is the probability that the market research firm would predict that it would be a success?

a) < 0.60
b) between 0.60 and 0.62
c) between 0.62 and 0.64
d) between 0.64 and 0.66
e) > 0.66

14) If the market research predicted that a product would be a success, what is the probability that it would actually be a success?

a) < 0.81
b) between 0.81 and 0.82
c) between 0.82 and 0.83
d) between 0.83 and 0.84
e) between 0.84 and 0.85

15) If the market research predicted that a product would not be a success, what is the probability that it would still be a success?

a) between 0.30 and 0.32
b) between 0.32 and 0.34
c) between 0.34 and 0.36
d) between 0.36 and 0.38
e) > 0.40

16) Let the two events, (S): “A new product is successful” and (N): “A new product is predicted to be a failure”. Which of the following statements is true?

a) S and N are independent
b) S and N are mutually exclusive
c) S and N are complimentary
d) S and N are collectively exhaustive
e) None of the above is true

Question 17-19 Alsace: The main beer-producing region of France – Part 1

Since the mid 19th century in Alsace, specialised breweries have taken over small family businesses and have introduced major innovations to control ale quality. Water drawn from the groundwater, barley, malt, yeast and hops are the ingredients of Alsace ale which gives it its unique density and alcohol content. We can assume that the density of Alsace ales follows a normal distribution with an average density of 19.5 degrees Plato and a standard deviation of 0.5 degree Plato. (*Degree Plato is an empirically derived hydrometer scale that measures how much denser the liquid is compared to water*).

17) We choose a microbrewery in Alsace randomly, what is the probability that the density from its ale is higher than 19.75 degrees Plato?

a) < 0.2
b) between 0.2 and 0.4
c) between 0.4 and 0.6
d) between 0.6 and 0.8
e) between 0.8 and 1.0

18) What is the probability that the density from a randomly selected microbrewery is between 19 and 20 degrees Plato?

a) < 0.2
b) between 0.2 and 0.4
c) between 0.4 and 0.6
d) between 0.6 and 0.8
e) between 0.8 and 1.0

19) What is the probability that in a randomly selected sample of five microbreweries the average density is lower than 19.3 degrees Plato?

a) < 0.2
b) between 0.2 and 0.4
c) between 0.4 and 0.6
d) between 0.6 and 0.8
e) between 0.8 and 1.0

Question 20-22 Alsace: The main beer-producing region of France – Part 2

This is additional information to Part 1.

Franck Müller, owner of a microbrewery in Alsace, has been working on a new fermentation process to increase the average density of local ales. A higher average density could be desirable since it is an indicator of alcohol content, and a higher alcohol content could be financially rewarding.

Frank claims that his new process yields a higher average density and will perhaps change the standard deviation of ales produced in the region. Frank would like to support his claim and has randomly selected batches from 6 microbreweries that have recently adopted his new process. Densities in degree Plato from the 6 microbreweries are given below:

20.2   19.5   19.6   20.6   20.3   21.0

20) To test the owner`s claim, the appropriate distribution to use is given by:

a) Poisson distribution
b) Exponential distribution
c) Z - distribution
d) CLT distribution
e) None of the above

21) What is the standard deviation in degree Plato of the average density of the sample of the 6 microbreweries?

a) < 0.2
b) between 0.2 and 0.4
c) between 0.4 and 0.6
d) between 0.6 and 0.8
e) between 0.8 and 1.0

22) What can you conclude about the new fermentation process?

a) The new fermentation process does not yield a higher average density compared to local ales and the average sample density is representative of the region
b) The new fermentation process does not yield a higher average density compared to local ales and the average sample density is not representative of the region
c) The new fermentation process yields a higher average density compared to local ales and the average sample density is not representative of the region
d) The new fermentation process yields a lower average density compared to local ales and the average sample density is not representative of the region
e) It is not possible to determine since the sample size is too small

Questions 23 - 26: Arts Classes

A grade 2 arts teacher is keeping track of how many times students spill their paint water each week. Last school year, the average rate of occurrence of such incidents was 6.35 spills per week.

23) During that year, what was the probability of getting fewer than 3 paint spills in a given week?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.05
d) between 0.05 and 0.2
e) between 0.2 and 1.0

24) Last week, the teacher had a grade 3 class instead of his usual grade 2 class. The number of spills was 2. If the probability of spills was the same as that for his grade 2 class, what would the probability of 2 or fewer spills be?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.05
d) between 0.05 and 0.2
e) between 0.2 and 1.0

25) The teacher wonders if this means that grade 3 students are less likely to spill their paint water.

a) Yes, they are less likely, because 2 < 3
b) No, they are not less likely, because 2 is not much smaller than 3
c) Yes, because the probability is <5%
d) No, because the probability is >5%
e) No, because the probability is <5%

26) The teacher also keeps track of the location in the classroom where the spills happen. The probability that a spill happens in the front half of the room (as opposed to the back half) is 67%. What is the probability of exactly 2 spills happening in the back half if there are 5 spills altogether?

a) < .001
b) between 0.001 and 0.005
c) between 0.005 and 0.02
d) between 0.02 and 0.5
e) between 0.5 and 1.0

Questions 27-30 : Sprayfoam

Sprayfoam insulation used in residential buildings is applied in layers of two inches. The application process is difficult, as the material is sprayed on as a liquid and then foams up somewhat irregularly. Therefore, the thickness of each layer is not very precise. An excellent and experienced contractor can apply the foam so that the thickness does not vary much, e.g. with a target of two inches, the variance is ¼ inch^2.

27) How would you expect the thickness to be distributed?

a) Uniform distribution
b) Exponential distribution
c) Normal distribution
d) Sampling distribution
e) Probability distribution

28) What thickness range do you expect 95% of the sprayed area to fall into?

a) 1-3 inches
b) 1.5-2.5 inches
c) 0.5-3.5 inches
d) 1.25-2.75 inches
e) We can't tell

29) You test one spot and find it's 1.15 inches. What is the probability of this or a thinner spot occurring?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.05
d) between 0.05 and 0.2
e) between 0.2 and 1.0

30) Would you conclude that the contractor is not very good?

a) Yes, because 1.15 is less than 2 inches
b) Yes, because the probability in question 29 is > 0.05
c) No, because the probability in question 29 is > 0.05
d) Yes, because the probability in question 29 is < 0.05
e) No, because the probability in question 29 is < 0.05

Questions 31-35 : Tomato Seeds

A certain type of tomato seeds germinates 90% of the time. A backyard farmer planted 25 seeds.

31) What is the probability that exactly 21 seeds germinate?

a) < .001
b) between 0.001 and 0.01
c) between 0.01 and 0.05
d) between 0.05 and 0.2
e) between 0.2 and 1.0

32) What is the probability that fewer than 23 seeds germinate?

a) < 0.2
b) between 0.2 and 0.4
c) between 0.4 and 0.6
d) between 0.6 and 0.8
e) between 0.8 and 1.0

33) What is the expected number of seeds that germinate?

a) > 24
b) between 23 and 23.9
c) between 21 and 22.9
d) between 2 and 3.9
e) < 2

34) A hobby farmer plants 200 seeds. We would like to know the probability that greater than 179 and less than 186 seeds germinate. Can this question be answered using the normal distribution?

a) Yes, you can always use the normal distribution
b) Yes, np and nq are larger than 10
c) No, np and nq are larger than 10
d) No, the sample is not large enough
e) No, this is not a scenario for a normal distribution

35) Assume you could use the normal distribution. What x values would you use to calculate the z values?

a) x_1=180.5; x_2= 185.5
b) x_1=179.5; x_2= 185.5
c) x_1=180.5; x_2= 184.5
d) x_1=179.5; x_2= 184.5
e) There really is no way to do this with a normal distribution

Q36-46 *iShop2Home*

iShop2Home provides *web* and *smart-device* interface support for on-line shopping. The web-based (*X*) and device-based (*Y*) transactions (reported in 1000's per week) are the basis of *iShop2Home*'s (weekly) revenue.

The first set of questions concern the random variables *X* and *Y* (namely web and device based services respectively).

For the purposes of this question the joint probability model between *X* and *Y* will be summarized with the following tabulated empirical probability model.

```
     Y=17 Y=20 Y=22 Y=24
X=44 0.00 0.04 0.04 0.03
X=48 0.04 0.16 0.16 0.02
X=52 0.12 0.19 0.08 0.00
X=55 0.06 0.03 0.03 0.00
```

36). Based on this probability model please determine the marginal probability model for web-based (*X*) services. Identify the choice below, whose probabilities match the marginal model that you determined.

a) 0, 0.04, 0.12, 0.06
b) 0, 0.16, 0.08, 0
c) 0.03, 0.02, 0, 0
d) 0.11, 0.38, 0.39, 0.12
e) 0.0275, 0.095, 0.0975, 0.03

Based on the marginal probability model for *X,* please determine the following:

37). Expected value *E[X]* (units, *thousands of transactions/week*).

a) 50
b) 49.8
c) 45.6
d) 46
e) 66.3

38). Standard deviation *SD[X]* (units, *thousands of transactions/week*).

a) 4.15
b) 4.79
c) 3.17
d) 22.9
e) 99.3

39). Management want to know the probability that the total transaction volume (=*X*+Y) will exceed 70 (units, *thousands of transactions/week*). Select the closest answer.

a) 0.33
b) 0.65
c) 1
d) 0.77
e) 0.35

To prevent cascading errors, please ignore your preceding calculations. Instead use the parameters tabulated below, along with a correlation coefficient of negative 0.45.

```
      param
 rv  μ     σ
  X 60     3
  Y 30     2
```

iShop2Home currently earns \$200 (web-based) and \$100 (device-based) per thousand web- and device- based transactions, respectively. Please determine the following for *iShop2Home*'s total weekly revenue.

40). The expected value (units of *\$/week*)

a) 12000
b) 3000
c) 7500
d) 45000
e) 15000

41). The standard deviation (units of *\$/week*)

a) 632
b) 540
c) 189737
d) 3606
e) 162111

42). What difference did the correlation coefficient make? How would your answers to the two previous questions have changed if the correlation was in fact zero?

a) Both the expected value and standard deviation would decrease
b) Both the expected value and standard deviation would increase
c) The expected value would increase while the standard deviation would be unchanged
d) The expected value would be unchanged while the standard deviation would increase
e) The expected value would be unchanged while the standard deviation would decrease

How would the following summaries of the total weekly revenue change if the price per unit of both *X* and *Y* were increased by 20 percent ($price_{new}$=price*1.2).

43). The coefficient of variation

a) increased by 20%
b) increased by 44%
c) increased by 10%
d) unchanged
e) decreased by 20%

44). The expected value

a) increased by 20%
b) increased by 44%
c) increased by 10%
d) unchanged
e) decreased by 20%

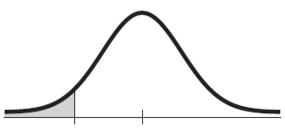
45). The variance

a) increased by 20%
b) increased by 44%
c) increased by 10%
d) unchanged
e) decreased by 20%

46). How would the correlation coefficient between *X* and *Y* change if 2000 **new** transactions/week were secured for **both** wired and devices (i.e.,$X_{new}=X+2$ and $Y_{new}=Y+2$).

a) It would be 2 times higher
b) Can not be determined
c) It would be 1.2 times higher
d) It would change sign
e) It would remain unchanged

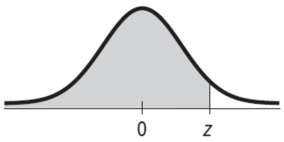
Areas under the standard Normal curve



Second decimal place in z										
0.09	*0.08*	*0.07*	*0.06*	*0.05*	*0.04*	*0.03*	*0.02*	*0.01*	*0.00*	*z*
									0.0000†	–3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	–3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	–3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	–3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	–3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	–3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	–3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	–3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	–3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	–3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	–2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026	–2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	–2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	–2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	–2.5
0.0064	0.0066	0.0068	0.0069	0.0071	00073	0.0075	0.0078	0.0080	0.0082	–2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	–2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	–2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	–2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	–2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	–1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	–1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	–1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	–1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	–1.5
0.0681	0.0694	0.0708	0.0721	0.07.35	0.0749	0.0764	0.0778	0.0793	0.0808	–1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	–1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0,1151	–1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	–1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	–1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	–0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	–0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	–0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	–0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	–0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446	–0.4
03483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	–0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	–0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	–0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	–0.0

†For $z \le -3.90$, the areas are 0.0000 to four decimal places.

Areas under the standard Normal curve

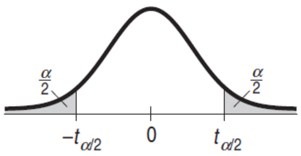


	Second decimal place in z									
z	*0.00*	*0.01*	*0.02*	*0.03*	*0.04*	*0.05*	*0.06*	*0.07*	*0.08*	*0.09*
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6703	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000†									

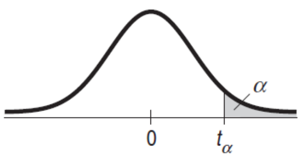
†For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

Table T

Values of t_α



Two tails



One tail

Two-tail probability	0.20	0.10	0.05	0.02	0.01	
One-tail probability	0.10	0.05	0.025	0.01	0.005	
df						df
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
30	1.310	1.697	2.042	2.457	2.750	30
32	1.309	1.694	2.037	2.449	2.738	32
35	1.306	1.690	2.030	2.438	2.725	35
40	1.303	1.684	2.021	2.423	2.704	40
45	1.301	1.679	2.014	2.412	2.690	45
50	1.299	3.676	2.009	2.403	2.678	50
60	1.296	1.671	2.000	2.390	2.660	60
75	1.293	1.665	1.992	2.377	2.643	75
100	1.290	1.660	1.984	2.364	2.626	100
120	1.289	1.658	1.980	2.358	2.617	120
140	1.288	1.656	1.977	2.353	2.611	140
180	1.286	1.653	1.973	2.347	2.603	180
250	1.285	1.651	1.969	2.341	2.596	250

ADM2303 formula sheet - Fall 2016

Probability theory

Rule of sum of probabilities:

$$P(S) = 1$$

Subtraction rule
(Let A^c be complement of A, i.e., Not A):

$$P(A) = 1 - P(A^c)$$

Addition rule for two mutually exclusive events (where $\cup$ connotes "or" aka union):

$$P(A \cup B) = P(A) + P(B)$$

Addition rule for two not mutually exclusive events:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Multiplication rule for two independent events (where $\cap$ connotes "and" aka intersection):

$$P(A \cap B) = P(A) \times P(B)$$

Multiplication rule for n independent events:

$$P(A_1 \cap A_2 \cap \ldots A_n) = P(A_1) \times P(A_2) \times \ldots \times P(A_n)$$

Multiplication rule for dependent events:

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

Partition rule: for a partition $B_1, B_2, \ldots, B_k$:

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

Bayes' formula:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

Events A and B are independent if:

$$P(A|B) = P(A) \text{ and } P(B|A) = P(B)$$
$$\text{or:} P(A \cap B) = P(A) \times P(B)$$

Random variables (RV)

Expected value of discrete RV X :

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(X = x_i)$$

Variance of discrete RV X: $\text{Var}(X) = \sigma^2$

$$= \sum_{i=1}^{n} (x_i - \mu)^2 P(X = x_i) \quad = \sum_{i=1}^{n} x_i^2 P(X = x_i) - \mu^2$$

Standard deviation of discrete RV X :

$$SD(X) = \sigma = \sqrt{Var(X)}$$

Coefficient of variation of discrete RV X :

$$CV(X) = \frac{SD(X)}{E(X)}$$

Correlation of two discrete RV X and Y : $Corr(X,Y)$

$$= \frac{\sum_{i=1}^{n} (x_i - \mu_x)(y_i - \mu_y))P(X = x_i \cap Y = y_i)}{s_x s_y}$$

Combining random variables

Adding a constant c to random variable X:

$$E(X \pm c) = E(X) \pm c$$
$$Var(X \pm c) = Var(X)$$

Multiplying random variable X by a constant a :

$$E(aX) = aE(X)$$
$$Var(aX) = a^2 Var(X)$$

Expected value of linear combination of RVs:

$$E(aX \pm bY \pm c) = aE(X) \pm bE(Y) \pm c$$

Variance of linear combination of **independent** RVs:

$$Var(aX \pm bY \pm c) = a^2 Var(X) + b^2 Var(Y)$$

Variance of linear combination of **dependent** RVs:

$$Var(aX \pm bY \pm c)$$
$$= a^2 Var(X) + b^2 Var(Y) \pm 2ab\,r\,\sigma(X)\sigma(Y)$$
$$= a^2 Var(X) + b^2 Var(Y) \pm 2ab\,Cov(X,Y)$$

Discrete and continuous distributions

The Binomial probability distribution:

$$P(X = x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, \ldots, n$$
$$E(X) = np, Var(X) = np(1-p)$$

The Poisson probability distribution
(If approx'n of binomial, $\lambda = np$):

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, \ldots$$
$$E(X) = \lambda, Var(X) = \lambda, e = 2.718$$

The Geometric probability distribution:

$$P(X = x) = (1-p)^{x-1} p \quad \text{for } x = 1, 2, \ldots$$
$$E(X) = \frac{1}{p}, Var(X) = \frac{1-p}{p^2}$$

The Normal distribution:

$$X \sim N(\mu, \sigma) \Rightarrow f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\right)$$
$$Z = \frac{X-\mu}{\sigma} \sim N(0,1) \Rightarrow P(Z < z) = \text{using normal table}$$

The Exponential distribution :

$$X \sim Expo(\lambda) \Rightarrow f(x) = \lambda e^{-\lambda x}$$
$$P(X \leq a) = 1 - e^{-a\lambda}$$
$$E(X) = \frac{1}{\lambda}, Var(X) = (\frac{1}{\lambda})^2$$

The Uniform distribution:

$$X \sim \text{Uniform}(a, b) \Rightarrow f(x) = \frac{1}{b-a} \quad \text{for } a \leq x \leq b$$
$$P(x_1 < X < x_2) = \frac{x_2 - x_1}{b-a}$$
$$E(X) = \frac{a+b}{2}, Var(X) = \frac{(b-a)^2}{12}$$

Descriptive statistics

Sample mean:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Sample variance:

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

Sample coefficient of variation:

$$CV = \frac{s}{\bar{x}}$$

Sample covariance:

$$\text{Cov}(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right)$$

Sample correlation:

$$r = \frac{Cov(x,y)}{s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}$$

Percentile:

- sort your data first

$$k_{th} \text{ percentile index} : i = (\frac{k}{100})(n+1)$$

- If i is integer, k_{th} percentile is the i_{th} value
- If i is not integer, k_{th} percentile is mean of the observations on either side of i

Boxplot elements:

$$IQR = Q_3 - Q_1$$
$$UpperLimit = Q_3 + 1.5(IQR)$$
$$LowerLimit = Q_1 - 1.5(IQR)$$

Normal approximation to Binomial

If $X \sim \text{Binomial}(n, p)$

If n is large i.e. $np \geq 10$ and $n(1-p) \geq 10$

$$\Rightarrow X \sim N\left(\mu_x = np, \sigma_x = \sqrt{np(1-p)}\right)$$

Sampling distributions for proportion

$$X \sim \text{Binomial}(n, p), \text{and } \hat{p} = \frac{X}{n}$$

If n is large i.e. $np \geq 10$ and $n(1-p) \geq 10$

$$\Rightarrow \hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Sampling distributions for mean

$$\text{If } X \sim N(\mu, \sigma) \Rightarrow \bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$
$$\text{If } X \sim N(\mu, \text{unknown}) \Rightarrow \bar{X} \sim t_{df=n-1}(\mu, \frac{s}{\sqrt{n}})$$
$$\text{If } X \sim \text{unknown}(\mu, \sigma) \xrightarrow[CLT]{\text{If n is large}} \bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$
$$\text{If } X \sim \text{unknown}(\mu, \text{unknown}) \xrightarrow[CLT]{\text{If n is large}} \bar{X} \sim t_{df=n-1}(\mu, \frac{s}{\sqrt{n}})$$

Finite population correction factor

In case of a finite population where $\frac{n}{N} > 10\%$ use :

$$\text{standard deviation} \times \sqrt{\frac{N-n}{N-1}}$$