MAT 1372 Statistics with Probability Practice Final Exam Spring 2018

**You may prepare a 2 page, one sheet, handwritten formula sheet.**

*Exam will be done in 2 parts. Once you have submitted part 1, you may work on part 2.*

**Part1: Excel or a graphing calculator if needed.**

Problems similar to 3 of the following 5 problems will appear (20 pts for each problem).

1. A fair coin is tossed (0 tail, 1 head) and a fair die (1,2,3,4) is rolled.
   1. List the sample space as a set of coordinates: (coin, die)

**S={(0,1),(0,2),(0,3),(0,4),(1,1),(1,2),(1,3),(1,4)}**

* 1. Write the event A= “sum of coordinates is 2 or 3” as set and find P(A).

**A={(0,2),(0,3),(1,1),(1,2)}, P(A)=.5**

* 1. Write the event B= “1st coordinate is 0” as a set and find P(B).

**B={(0,1),(0,2),(0,3),(0,4)}, P(B)=.5**

* 1. Write the event A∩B as a set and find P(A∩B).

**A∩B={(0,2),(0,3)}, P(A and B)=.25**

* 1. Are A and B independent?

**Yes, P(A)\*P(B)=.25=P(A∩B)**

1. The distribution of grades on a quiz taken

by 100 students is given in the table:



1. Compute the mean and standard deviation of this distribution.

**Mean is 2.54 and std dev is 0.93.**

(b) Find mean & std err (std dev for sample mean) for random sample of size n=16.

**Sample mean’s mean stays the same (2.54).**

**The std err (std dev for sample mean) = .93/sqrt(16)=.23**

(c) What value of n is needed so that std err (std dev for sample mean) ≤ .1? (round up)

**.93/sqrt(n)=.1, solving for n gives sqrt(n)=.93/.1 or n=(9.3)^2=86.5 or 87 (n is integer).**

1. 4% of a clinic’s patients are known to have Lyme’s disease. A test is developed that is positive in 98% of patients with Lyme’s disease, but it is also positive in 3% of patients who do not have disease. Fill in the following table (use all available digits).





1. What is probability that test comes out positive for Lyme’s disease?

**.04\*.98+.96\*.03=.068**

1. What is probability that a person has Lyme’s disease given positive test?

**P(Yes|positive)=P(Yes∩positive)/ P(positive)=.0392/.068=.576 or about 58%.**

In other words, there is a 42% chance of a **false positive**.

1. What is probability that a person has disease given a negative test (false negative)?

**P(Yes|negative)=P(Yes∩ negative)/ P(negative)=.0008/.932= .0009 or about .1%**

This is typical; oftentimes false positives are more common than false negatives.

1. A manufacturer wants to know if an experimental version of its child toothpaste differs in its ability to prevent tooth decay from the existing formula. For children using its existing formula, cavities per yr. are normal with mean 3.03 & standard deviation 1.12. A study of 100 children using new version found average of 2.83 cavities. **Two tailed**.
2. Calculate the SE (standard error, the standard deviation for the sample mean).

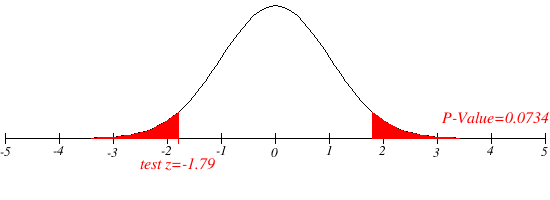
**H0: µ=3.03**

**H1: µ≠3.03**

**α=5% (default)**

**SE= σ/sqrt(n)= 1.12/sqrt(100)= .112**

1. Draw the normal curve and label sample mean, mean plus/minus SE, mean plus/minus 2SE and the average experimental value. Sketch in appropriate tail.



2.81 2.92 3.03 3.14 3.25

1. Standardize and mark off the z-value on the Z labeling of the horizontal axis in b).

**(2.83 – 3.03)/(.11)= -.20/.112=-1.79**

1. Use your calculator **2\*Normalcdf(-1E99,-1.79)** or Excel **2\*NORM.S.DIST(-1.79)** to find the p-value **=2\*3.7%=7.4%**
2. Does the experimental version’s cavity prevention differ from the existing formula?

**Since the pvalue is more than α=5%, we do NOT reject H0. In other words, we cannot say that the cavity prevention of the experimental version differs from that of the existing formula.**

1. Mauna Loa is a major Hawaiian volcano. Its eruptions have been shown to follow roughly a Poisson process. Within a 10-year period, on average there are 3 eruptions.

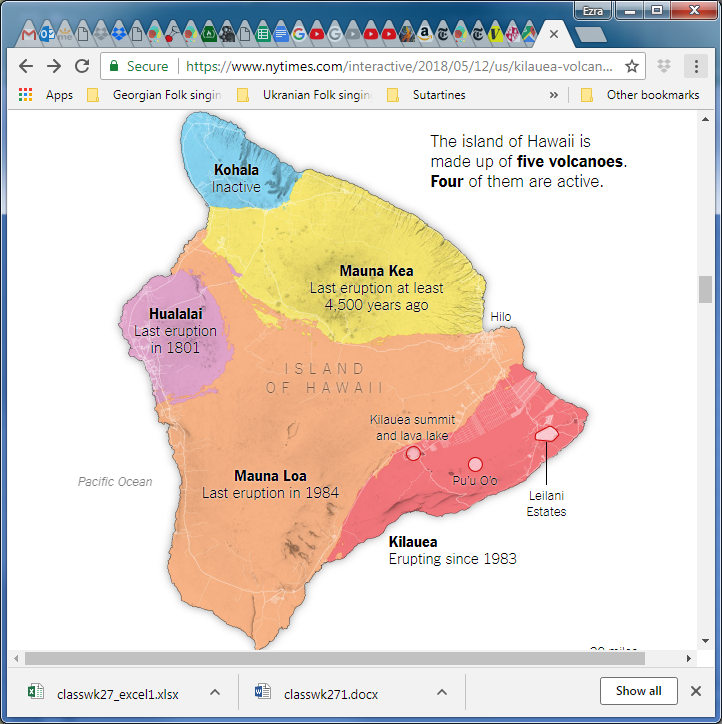
BE PREPARED TO USE EITHER or BOTH of definition/built-in calculator/Excel functions.

* 1. In the next 10 years, what is the chance that there will be at least one eruption?

**λ=3, Using def: 1- exp(-3), using Excel: 1-poisson.dist(3,0,false)=95.0%**

* 1. In the next 10 years, what is the chance that there will be exactly 3 eruptions?

λ**=3, Using def: exp(-3)\*3^2/6, using Excel: poisson.dist(3,3,false)=22.4%**

* 1. In the next 10 years, what is the chance that there will be fewer than 3 eruptions?

λ**=3, using Excel: poisson.dist(2,3,true)=42.3%**

* 1. In the next 20 years, what is the chance that there will be exactly 6 eruptions?

λ**=6, using Excel: poisson.dist(6,6,false)=16.1%**

* 1. Why or why not is your answer to d. the square of your answer to b. (ans b^2=5.0%)? **It is not because the 6 eruptions could be distributed in many different ways between to 2 groups of 3 years, not just 3 and 3.**

Picture comes from [NYT\_art](https://www.nytimes.com/interactive/2018/05/12/us/kilauea-volcano-lava-leilani-estates-hawaii.html)

See also [Science\_news\_art](https://www.sciencenews.org/article/how-long-volcano-kilauea-hawaii-eruption-last?)

**Answers to Part II (problems 5 and 6) can be found in the Excel file, except for the graphs, those should be put into your blue book.**

**Part 2: Down- and up-load the Excel file from Blackboard.**

**Please make sure your name becomes part of the file name.**

**NO PICTURE NO CREDIT! (Put graphs into your blue book)**

Problems similar to each of the following problems will appear (20 pts each).

1. The FDA has decided to measure 8-ounce cans of Hershey Cocoa. The weights of 50 containers are found in the accompanying excel file. **Single tailed.**
   1. Use the data to calculate the sample standard deviation and mean.
   2. Find the SE (standard error, the standard deviation for the sample mean).
   3. Use an appropriate test and the p-value approach to determine if the cans do not contain the claimed weight.
2. Do people spend more time reading ads that agree with their disposition? An antismoking group conducted a survey asking those who saw an advertisement against smoking if (1) they read headline only, (2) some detail, or (3) most of advertisement. The questionnaire asks respondents to identify themselves as heavy smokers (1), moderate smokers (2), light smokers (3) or nonsmokers (4). The result is found in the accompanying excel file (col A: type of smoker, col B: amt of advertisement read). Does how much you smoke affect how much you read of an antismoking advertisement? Use a pvalue approach at the 10% level.