Name ____

Fall 2024 **STAT 5870-1/A** **Exam I**

(100 points)

Instructions:

- Write your name on the top, but do not open the exam.
- You are allowed to use one 8.5" x 11" page of notes (front and back) and a calculator.
- A total of 4 pages with a front and back.
- For full/partial credit, show all your work.

1. Identify the most appropriate distribution to model the data described. For each question, **circle one** of binomial, Poisson, or normal. (2 pts each)

   (a) The height of a tree in Sycamore row

   binomial　　Poisson　　normal

   Answer: normal

   (b) Number of musical notes played incorrectly in a measure with 9 notes

   binomial　　Poisson　　normal

   Answer: binomial

   (c) Number of coding bugs in the R package 'emmeans'

   binomial　　Poisson　　normal

   Answer: Poisson

   (d) Time taken to complete a 5k

   binomial　　Poisson　　normal

   Answer: normal

   (e) Number of songs Taylor Swift will sing off her Evermore album at her next concert

   binomial　　Poisson　　normal

   Answer: binomial

   (f) Amount of money donated to Story Theatre Company this year

   binomial　　Poisson　　normal

   Answer: normal

   (g) Score for an Olympic dive

   binomial　　Poisson　　normal

   Answer: normal

   (h) In a survey of 50 people, the number who use liquid detergent when washing clothes

   binomial　　Poisson　　normal

   Answer: binomial

   (i) Calories consumed in a day

   binomial　　Poisson　　normal

   Answer: normal

   (j) Number of times you would clap when Dr. Niemi wins the Iowa State University Teaching Excellence Award

   binomial　　Poisson　　normal

   Answer: Poisson

2. Consider the following probability density function (pdf)

$$f(x) = \begin{cases} (x-1)/2 & 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Explain why this is a valid pdf. (4 pts)

Answer: The function is non-negative, i.e. $f(x) \geq 0$ for all x. The integral (area under the curve) is 1 since the function forms a triangle with a base of 2 and a high of 1.

(b) State the image for a random variable with this probability density function. (4 pts)

Answer: The function is positive on the interval (1,3) and thus (1,3) is the image.

(c) Determine the cumulative distribution function for a random variable with this pdf. (8 pts)

Answer:

$$F(x) = \begin{cases} 0 & x < 1 \\ (x-1)^2/4 & 1 < x < 3 \\ 1 & x > 3 \end{cases}$$

Equalities can be placed anywhere.

(d) State the integral that would be used to find the mean for a random variable with this pdf. More points will be rewarded for an integral this is more specific, but you do NOT need to solve the integral for full credit. (4 pts)

Answer:

$$\int_{-\infty}^{\infty} x f(x) dx = \int_{1}^{3} x \frac{(x-1)}{2} dx$$

3. In 2022, a mother tested positive for opiates shortly after giving birth to her baby. Data from 2013 suggests that only 6 of every 1,000 pregnant mothers use opiates. For the test used, the sensitivity (the probability of testing positive if opiates have been used) is 95% while the specificity (the probability of testing negative if opiates have not been used) is 85%. Based on this positive test, what is the probability this mother did use opiates. (20 pts)

Answer: This problem is based on a real case and numbers were taken from relevant sources:

- `https://www.themarshallproject.org/2024/09/09/drug-test-pregnancy-pennsylvania-cal`
- `https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2017/08/opioid-use-and-opioid-use-disorder-in-pregnancy`
- `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6120972/#:~:text=(2012).,and%20cocaine%20despite%20manufacturer%20recommendations.`

Let

- $+$ $(-)$ indicate a positive (negative) test result and
- U (N) indicate a user (non-user).

We know

- $P(U) = 6/1000 = 0.006$
- $P(+|U) = 0.95$
- $P(-|N) = 0.85$

We use Bayes' Rule to find

$$\begin{aligned} P(U|+) &= \frac{P(+|U)P(U)}{P(+|U)P(U)+P(+|N)P(N)} \\ &= \frac{P(+|U)P(U)}{P(+|U)P(U)+[1-P(-|N)][1-P(U)]} \\ &= \frac{0.95\times 0.006}{0.95\times 0.006+[1-0.85][1-0.006]} \\ &= 0.0368 \end{aligned}$$

Thus the probability of an opiate user given the positive test result is only 3.68%.

4. Carbon has two stable, non-radioactive isotopes, ^{12}C and ^{13}C, with relative isotopic abundances of 98.89% and 1.11%, respectively. The molecular formula for cholesterol is $C_{27}H_{44}O$, i.e. there are 27 carbon molecules.

   Answer: This question taken from https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Analytical_Chemistry_2.1_(Harvey)/04%3A_Evaluating_Analytical_Data/4.04%3A_The_Distribution_of_Measurements_and_Results.

   Let X be the number of ^{13}C atoms in a molecule of cholesterol. Assume $X \sim Bin(n, p)$ with the following values for n and p:

```
n <- 27
p <- 0.0111
```

   (a) What is the mean number of ^{13}C atoms in a molecule of cholesterol? (4 pts)
   Answer:

```
n * p

## [1] 0.2997
```

   (b) What is the standard deviation for the number of ^{13}C atoms in a molecule of cholesterol? (4 pts)
   Answer:

```
sqrt(n * p * (1-p))

## [1] 0.5444018
```

   (c) What is the probability that a molecule of cholesterol has no ^{13}C atoms? (4 pts)
   Answer:

```
dbinom(0, size = n, prob = p) # OR

## [1] 0.7397997

(1 - p)^n

## [1] 0.7397997
```

   (d) What is the probability that a molecule of cholesterol has at least one atom of ^{13}C? (4 pts)
   Answer:

```
1 - dbinom(0, size = n, prob = p)

## [1] 0.2602003
```

(e) If we want to assure the probability of at least one atom of ^{13}C in a molecule of cholesterol is less than 0.05, what would the relative isotropic abundance of ^{13}C need to be? (4 pts)

Answer: We need

$$1 - (1-p)^{27} = 0.05 \quad \Longrightarrow \quad p = 1 - (1-0.05)^{1/27}$$

```
1 - (1 - .05)^(1/27)

## [1] 0.001897948
```

5. The Ames air quality monitor is set to collect $PM_{2.5}$ (particulate matter less than 2.5 μm) once per day and report the total amount of pollutant measured after 30 days. Answer:

```
n         <- 30
mu        <- 9.9
threshold <- 9
sigma     <- round( (mu - threshold) * sqrt(n), 1)
z         <- (threshold - mu) / (sigma / sqrt(n))
```

Currently, the average $PM_{2.5}$ measurement in Ames is 9.9 $\mu g/m^3$ per day with a standard deviation of 4.9 $\mu g/m^3$ per day.

(a) What is the expected sum of Ames $PM_{2.5}$ measurements over the next 30 days? (4 pts)
Answer:

```
n * mu

## [1] 297
```

(b) Assuming measurements are independent, what is the variance of the sum of Ames $PM_{2.5}$ measurements over the next 30 days? (4 pts)
Answer:

```
n * sigma^2

## [1] 720.3
```

As a side note, this independence assumption is probably not reasonable as the observations are correlated in time.

(c) What is the approximate probability the 30-day **average** is greater than the national standard of $9\,\mu g/m^3$? (8 pts)
Answer: Let $\overline{X}$ be the average $PM_{2.5}$ measurement in Ames over 30 days. From the above, we know $E[\overline{X}] = 9.9\,\mu g/m^3$ and $SD[\overline{X}] = 4.9/\sqrt{30}\,\mu g/m^3$. Assume CLT applies and thus $\overline{X} \sim N(9.9, 4.9^2/30)$.
Calculate

$$
\begin{aligned}
P(\overline{X} > 9) &= 1 - P(\overline{X} \leq 9) \\
&= 1 - P\left(\frac{\overline{X}-9.9}{4.9/\sqrt{30}} \leq \frac{9-9.9}{4.9/\sqrt{30}}\right) \\
&= 1 - P(Z \leq -1.006021) \\
&= 1 - [1 - P(Z \leq 1.006021)] = P(Z \leq 1.006021) \\
&= 0.8428
\end{aligned}
$$

(d) What would the expected daily Ames $PM_{2.5}$ measurement need to be so that the probability of the 30-day average being greater than $9\,\mu g/m^3$ is 16%? (4 pts)

Answer: The easy way to see the answer is to realize that $1 - 0.16 = 0.84$ which is the answer to the previous problem. For the previous problem, the average was 1 sd above the mean, now the average needs to be 1 sd below the mean. So, a mean daily measurement of $9 - 0.9 = 8.1\mu g/m^3$ would work.

Formally, $0.16 = 1 - P(Z < 1)$ and thus

$$1 = \frac{9 - \mu}{4.9/\sqrt{30}} \implies \mu = 9 - 4.9/\sqrt{30} = 8.1053865$$

Please use as scratch paper.

Table 1: Cumulative distribution function, $P(Z \leq z)$, for standard normal

z	0	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	0.6103	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