SP24 STAT340 Midterm in-class portion

First (given) name:

Write here: _____

Last (family) name:

Write here: _____

Lecture section:

Circle one: **Bi's section**   **Brian's section**   **Yongyi's section**

Rules:

- You must show work for all computations (unless otherwise specified) to receive full credit.
- You **do NOT need to simplify any expressions** you write down.
- Note some of the multiple choice are **choose ONE** and some are **choose ALL that apply**, please pay attention to the instruction and select the appropriate number of responses!

Points:

MC1-2 (/4)	MC3-7 (/10)	MC8-9 (/4)	SA1 (/4)	SA2 (/4)	SA3 (/4)	Total (/30)

Multiple Choice 2pts each

NOTE: For each multiple choice below, choose ONE means there is EXACTLY one right answer, choose ALL means there is AT LEAST one right answer.
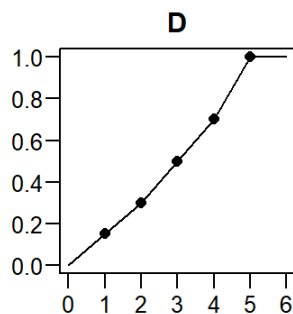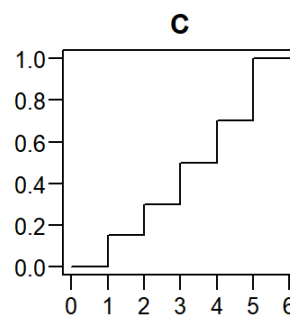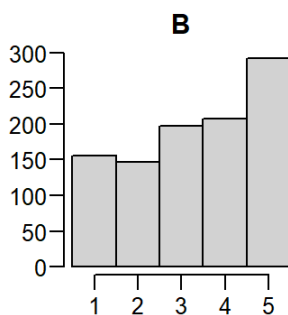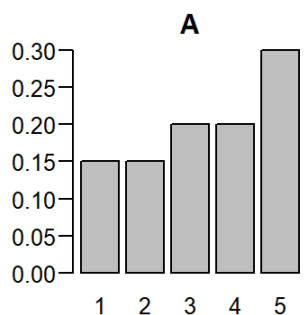
MC1

Which of the following random variables has the biggest mean μ? If there are multiple tied for largest, **choose ALL that apply!**

    a. Binomial(20,0.2)
    b. Normal(4,4²)
    c. Geometric(0.2)
    d. Poisson(4)
    e. Exponential(0.25)

MC2

I have a funky die: it is weighted and it has the value 5 twice and no 6 (i.e. the faces are 1,2,3,4,5,5). The probabilities $P(X = k)$ for $k = 1, 2, \ldots, 5$ are in order $0.15, 0.15, 0.20, 0.20, 0.30$. Suppose I roll the die once and X represents the value I roll. Which of the following is the CDF of X? **Choose ONE!**

MC3

If A, B are independent events, which of the following is true? **Choose ONE!**

    a. $P(A \cup B) = P(A)P(B)$
    b. $P(A \cap B) = P(A) + P(B)$
    c. $P(A \cap B) = P(A|B)$
    d. $P(A|B) = P(A)$
    e. $P(A|B) = P(B|A)$

MC4

Which of the following are true? **Choose ALL that apply!**

    a. If X is a continuous random variable, then for any k, $\mathbf{P}(X = k) = 0$.
    b. We cannot use a normal distribution to model human heights because a normal random variable can take negative values.
    c. If X, Y are two RVs that are not independent, then it's NOT ALWAYS true that $Var(X + Y) = Var(X) + Var(Y)$.
    d. If X, Y are two RVs that are not independent, then it's NOT ALWAYS true that $E(X + Y) = E(X) + E(Y)$.
    e. Two events A, B can be both independent and mutually exclusive.

MC5

Let $X \sim N(10, 2^2)$, which of the following is closest to $P(X < 12)$? Use the empirical rule. (Note that we use the notation N(μ,σ²) in STAT340). **Choose ONE!**

    a. 0.34
    b. 0.68
    c. 0.84
    d. 0.95
    e. 0.997

MC6

You are a prosecutor for the SEC (Securities and Exchange Commission) suing companies for fraudulent financial reporting. Suppose we have the usual null (i.e. no fraud exists). Match each of the options below with one of these test outcomes: TP, FP, TN, FN

    a. A fraudulent company is found guilty.   _____
    b. A fraudulent company is found not guilty. _____
    c. An innocent company is found guilty.   _____
    d. An innocent company is found not guilty. _____

MC7

Let T represent the theoretical distribution assuming the null hypothesis of some statistic t you have just computed. Furthermore, suppose T is symmetric around 0. Which of the following is NOT a valid p-value for some pair of null/alternative hypotheses you could test? **Choose ALL that apply!**

    a. $P(T \geq t)$
    b. $P(T \leq t)$
    c. $P(|T| \geq |t|)$
    d. $P(|T| \leq |t|)$
    e. They are all valid for some pair of null/alternative hypotheses

MC8-9

The following statement applies to both MC8 and MC9: Let $X_1, X_2, \ldots X_n$ be an i.i.d. (i.e. independent and identically distributed) sample where X_i has mean μ and variance σ^2.
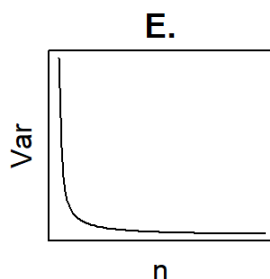
MC8

Which of the following shows the EXPECTED (i.e. typical) relationship between the sample variance S^2 and n? Recall that $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$. **Choose ONE!**
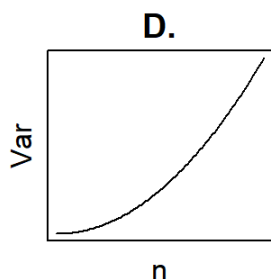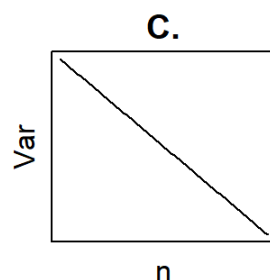
A.

B.

C.

D.

E.
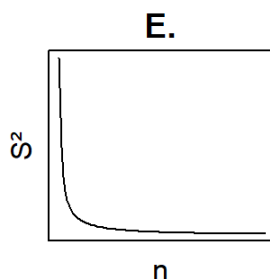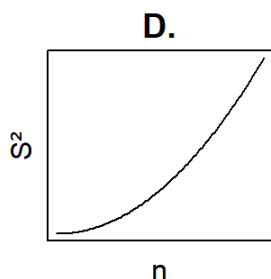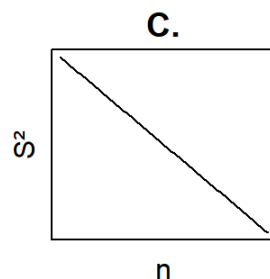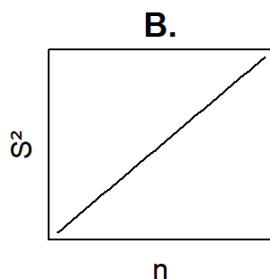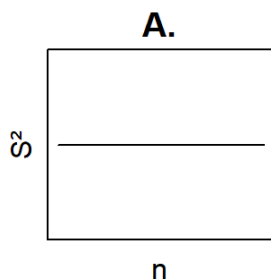
MC9

Which of the following shows the EXPECTED (i.e. typical) relationship between the mean variance $Var(\bar{X})$ and n? Recall $Var(\bar{X}) = Var(\frac{1}{n} \sum_{i=1}^{n} X_i)$. **Choose ONE!**

A.

B.

C.

D.

E.
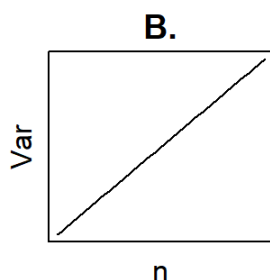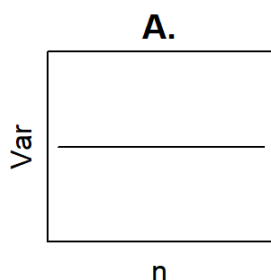
Short Answer 4pts each

SA1

Suppose you are going to spend the afternoon in the UW Arboretum doing some bird watching. Describe a random variable related to your bird-watching afternoon that could be modeled using each of the following. Make sure to state and justify any assumptions you make!

  a. A uniform distribution
  b. A binomial distribution
  c. A geometric distribution
  d. A normal distribution

SA2

In a bag of 48 rainbow candies I get the following:

```
##      red orange yellow  green   blue purple
##       11      8      5      8      4     12
```

a. I really like blue and orange. If I eat a candy at random and I really like it, what is the probability that it is orange?

b. I decide to grab and eat candies at random until I get one that I really like. I decide to model the number of candies I get BEFORE getting one I really like as a geometric random variable with $p = 12/48$. Is this a good model? Why or why not?

c. Define the colors blue and purple to be "dark" colors, and the rest to be non-dark. Is getting a dark candy mutually exclusive with getting a color I really like? Explain why or why not.

d. Is getting a dark candy independent from getting a color I really like? Explain why or why not.

SA3

A medical supplier Trimax wants to build a better rapid COVID antigen test. They have three models: Trimax-A, Trimax-B and Trimax-C. They take 50 individuals who are confirmed to have COVID-19 and 50 individuals who are confirmed to not have COVID-19 and ask them to take all three tests. The null hypothesis in each test is that the person does not have COVID. A "FALSE" result is a negative result (i.e. no COVID). Here are the results.

```
## , , Trimax = A          , , Trimax = B          , , Trimax = C
##        Result                   Result                  Result
## Covid   FALSE TRUE      Covid   FALSE TRUE      Covid   FALSE TRUE
##   FALSE    41    9        FALSE    47    3        FALSE    39    11
##   TRUE      4   46        TRUE      7   43        TRUE      0    50
```

    a. Calculate the Type 1 error rate of Trimax-A

    b. Calculate the power of Trimax-B

    c. Calculate the accuracy (i.e. proportion of all results that are correct) of Trimax-C.

    d. Which test do you think they should market, and why?