STAT 511 fa 2019 Exam I

Karl B. Gregory

- Do not open this test until told to do so.
- No calculators allowed; no notes allowed; no books allowed.
- Simplify all answers.
- SHOW YOUR WORK so that PARTIAL CREDIT may be given.

Chebychev's inequality: For any random variable X with mean μ_X and variance σ_X^2 and any constant $K > 0$, we have

$$P_X(|X - \mu_X| < K\sigma_X) \geq 1 - \frac{1}{K^2}.$$

1. Consider rolling a 6-sided die with sides ⚀, ⚁, ⚂, ⚃, ⚄, and ⚅ and define the random variable

$$X(s) = \begin{cases} 1 & \text{if } s \in \{⚀, ⚁, ⚂\} \\ 2 & \text{if } s \in \{⚃, ⚄\} \\ 3 & \text{if } s \in \{⚅\}. \end{cases}$$

(a) Give the support of X.

Solution: The random variable X can take any of the values in $\mathcal{X} = \{1, 2, 3\}$.

(b) Tabulate the probability distribution of X with a table of the form

x	$\cdots$
$P_X(X = x)$	$\cdots$

Solution: We have

x	1	2	3
$P_X(X = x)$	3/6	2/6	1/6

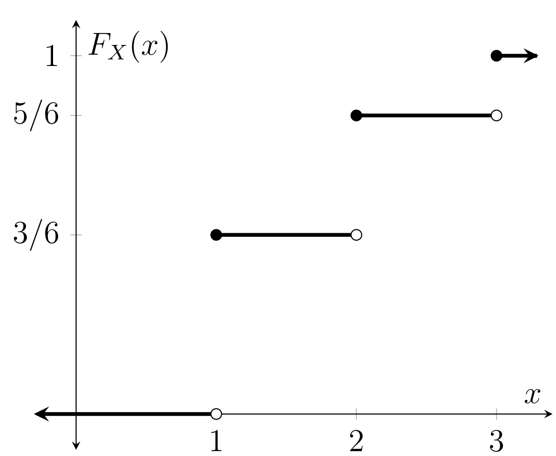
(c) Write down the cdf F_X of X, making sure to define $F_X(x)$ for all $x \in \mathbb{R}$.

Solution: The cdf F_X of X is the function given by

$$F_X(x) = \begin{cases} 0, & -\infty < x < 1 \\ 3/6, & 1 \le x < 2 \\ 5/6, & 2 \le x < 3 \\ 1, & 3 \le x < \infty \end{cases}$$

(d) Draw a detailed picture of the cdf F_X.

Solution:



(e) Give the following probabilities:

i. $P_X(X \leq 1/2)$

> **Solution:** This is 0, since $F_X(1/2) = 0$, which makes sense because X cannot take any values this small.

ii. $P_X(X \leq 2.5)$

> **Solution:** This is given by $F_X(2.5) = 5/6$.

iii. $P_X(1 < X \leq 3)$

> **Solution:** This is given by $F_X(3) - F_X(1) = 3/6$.

(f) Compute the expected value $\mathbb{E}X$ of X.

> **Solution:** The expected value is given by
>
> $$\mathbb{E}X = 1(3/6) + 2(2/6) + 3(1/6) = (1 + 4 + 3)/6 = 5/3.$$

(g) Compute the variance $\operatorname{Var} X$ of X.

> **Solution:** Use $\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2$. We have
>
> $$\mathbb{E}X^2 = 1^2(3/6) + 2^2(2/6) + 3^2(1/6) = (3 + 8 + 9)/6 = 10/3,$$
>
> so
>
> $$\operatorname{Var} X = 10/3 - (5/3)^2 = (30 - 25)/9 = 5/9.$$

(h) Use Chebychev's inequality to give an interval within which X will fall with probability at least $1 - 1/16 = 0.9375$.

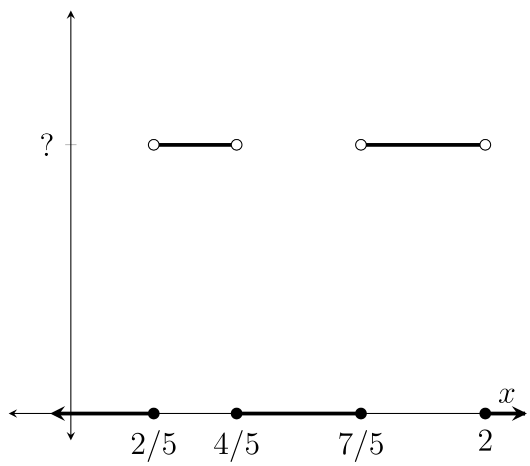
> **Solution:** The standard deviation is X is $\sigma_X = \sqrt{5}/3$. According to Chebychev's inequality, X will fall within 4 standard deviations of its mean with probability at least $1 - 1/16$. The corresponding interval is given by
>
> $$(5/3 - 4\sqrt{5}/3, 5/3 + 4\sqrt{5}/3).$$

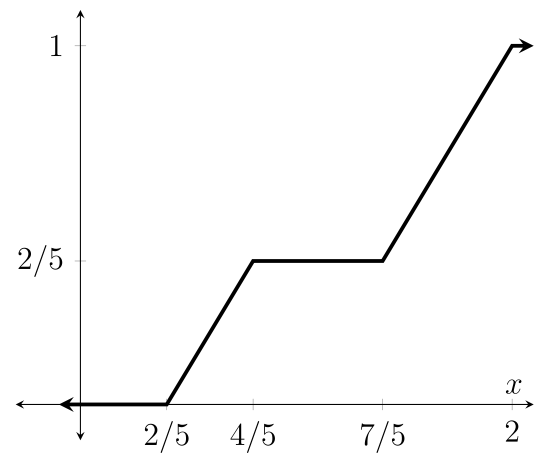
(i) Comment on whether you think the interval you gave in part (h) is useful for this random variable.

> **Solution:** We find that the entire support of X fits into this interval. That is
>
> $$\{1, 2, 3\} \subset (5/3 - 4\sqrt{5}/3, 5/3 + 4\sqrt{5}/3)$$

since $5/3 - 4\sqrt{2}/3 < 1$ and $5/3 + 4\sqrt{2}/3 > 3$. So the interval does not give us any extra information about what values X is most likely to take (Chebychev's inequality certainly *is* useful in many cases, just not in this case).

2. One of the two plots below shows the cdf of a random variable X and the other shows the pdf of the same random variable.





(a) Which plot shows the cdf?

Solution: The right-hand plot shows the cdf because the function is non-decreasing, "begins" at 0 and "ends" at 1, and it is right-continuous (it is moreover continuous).

(b) Is the random variable X discrete or continuous?

Solution: The random variable X is a continuous random variable because the cdf is continuous.

(c) Give the support of X.

Solution: The pdf, given in the left-hand plot, is positive over the set of values

$$\mathcal{X} = (2/5, 4/5) \cup (7/5, 2),$$

so this is the support.

(d) Give the following probabilities:

i. $P_X(X \leq 1)$

Solution: The probability is given by $F_X(1) = 2/5$, which we see from the right-hand plot.

ii. $P_X(X = 7/5)$

> **Solution:** This is 0 because the random variable is continuous.

iii. $P_X(4/5 < X < 2)$

> **Solution:** The probability is given by $F_X(2) - F_X(4/5) = 3/5$.

(e) Give the height of the function in the left-hand plot over the intervals $(2/5, 4/5)$ and $(7/5, 2)$.

> **Solution:** Since the slope of the cdf F_X over these intervals is 1, the height of the pdf over these intervals must be equal to 1.

3. (a) Give the number of unique sequences of letters that can be created with the letters in *borogoves*. You do not need to simplify your answer.

> **Solution:** Ah, *borogoves*! As in *all mimsy were the borogoves* from "Jabberwocky" by the mathematician-poet Lewis Carroll. Now let's see, the 9 letters can be arranged in 9! ways, but the *o*'s can be ordered amongst themselves in 3! ways, so the total number of unique sequences is
>
> $$\frac{9!}{3!}.$$

(b) Consider the following set of words:

jaws the that claws the catch bite that

i. Suppose you draw two words without replacement from the above set of words. Give the probability that you draw the words *claws* and *the*. The order in which you draw them does not matter. Simplify your answer.

> **Solution:** There are $\binom{8}{2}$ ways to draw 2 of the 8 words when order does not matter. There is $\binom{1}{1} = 1$ to draw the word *claws* and there are $\binom{2}{1}$ ways to draw the word *the*. So the probability is
>
> $$\frac{\binom{1}{1}\binom{2}{1}}{\binom{8}{2}} = 1/14.$$
>
> To look at it another way, we could first draw *claws* and then drawn *the*, for which the probability is $(1/8)(2/7)$, or we could first draw *the* and then draw *claws*, for which the probability is $(2/8)(1/7)$. The sum of these probabilities is
>
> $$(1/8)(2/7) + (2/8)(1/7) = 4/56 = 1/14.$$

ii. Suppose you draw one word at a time from the above set of words until you have drawn all the words. Give the probability that your sequence of draws results in the phrase *the jaws that bite the claws that catch*. You do not have to simplify your answer.

Solution: We can draw all 8 of the 8 words in 8! ways; 4 of these ways will result in the phrase *the jaws that bite the claws that catch*, since we can transpose the two *the*'s and the two *that*'s. So the probability is

$$\frac{4}{8!}.$$

4. Suppose 1/10 of all the text messages you receive come from family members, and 1/5 of the messages from family members come before 8:00 am. In addition, suppose that 19/20 of the messages you receive from non-family members come after 8:00 am.

   (a) What is the proportion of text messages you receive before 8:00 am?

Solution: Let F be the event that a message is from a family member and let B be the event that it comes before 8:00 am. We wish to find $P(B)$. We have

$$\begin{aligned} P(B) &= P(B \cap F) + P(B \cap F^C) \\ &= P(B|F)P(F) + P(B|F^C)P(F^C) \\ &= P(B|F)P(F) + [1 - P(B^C|F^C)][1 - P(F)] \\ &= (1/5)(1/10) + (1 - 19/20)(1 - 1/10) \\ &= 13/200 \end{aligned}$$

   (b) If you receive a text message before 8:00 am, what is the probability that it is from a family member?

Solution: We wish to find $P(F|B)$, which is given by

$$P(F|B) = \frac{P(F \cap B)}{P(B)} = \frac{P(B|F)P(F)}{P(B)} = \frac{(1/5)(1/10)}{13/200} = 4/13.$$