Chebyshev's Inequality: Let X be a discrete random variable with mean μ and standard deviation σ. For any $k \geq 1$,
$P(|X - \mu| \geq k\sigma) \leq \dfrac{1}{k^2}$.
In words, the probability that X is at least k standard deviations away from its mean is at most $\frac{1}{k^2}$.

1. Verify that Chebyshev's Inequality holds with $k = 2$ for the random variable X from Problem 4 from the previous class, using the value $k = 2$. That is, check that $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$.

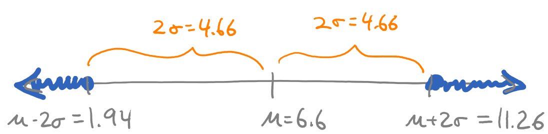
Recall $\mu_X = 6.6$ and $\sigma_X = \sqrt{5.44} \approx 2.33$

Consider: $P(|X - 6.6| \geq 2(2.33))$

$= P(|X - 6.6| \geq 4.66)$

$= P(X \leq 1.94 \text{ or } X \geq 11.26)$

$= 0$ (Since X takes no values less than 4 or greater than 10.)



Since $P(|X - \mu| \geq 2\sigma) = 0$, which is less than $\frac{1}{2^2} = \frac{1}{k^2}$,

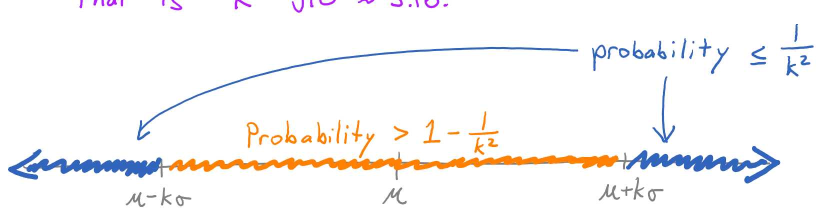
we see that Chebyshev's inequality holds in this case.

2. The number of equipment breakdowns in a manufacturing plant averages 4 per week, with standard deviation 0.7 per week.

(a) Find an interval that includes at least 90% of the weekly figures for the number of breakdowns.

Apply Chebyshev's Inequality with k that solves $\frac{1}{k^2} = 0.1$.

That is $k = \sqrt{10} \approx 3.16$.



We want:
$1 - \frac{1}{k^2} \geq 0.9$

Chebyshev's Inequality then says:

$P(|X-4| \geq 3.16(0.7)) = P(|X-4| \geq 2.21)$
$= P(X \leq 1.79 \text{ or } X \geq 6.21) \leq 0.1$

Take the complement to flip the inequality:
$P(1.79 < X < 6.21) > 0.9$

So the interval $(1.79, 6.21)$ includes at least 90% of the numbers of weekly breakdowns.

(b) A plant supervisor promises that the number of breakdowns will rarely exceed 7 in a one-week period. Is the supervisor justified in making this claim? Why?

From part (a), we see that 90% of weeks have less than 7 breakdowns.

We can do even better if we apply Chebyshev's Inequality with $k=5$:
$P(|X-4| \geq 5(0.7)) = P(X \leq 0.5 \text{ or } X \geq 7.5) = P(X=0) + P(X>7) \leq \frac{1}{5^2}$

So $P(X>7) \leq \frac{1}{25} = 0.04$.

Thus, the probability of more than 7 breakdowns in a week is not greater than 0.04. The supervisor's claim seems justified.

3. Suppose that 45% of the phone calls you receive are scam calls. Assume that the probability of a scam call is independent from one call to the next.

(a) Let $X = 1$ if the next call you receive is from a scam call, and $X = 0$ otherwise. What type of random variable is X? What are its mean and standard deviation?

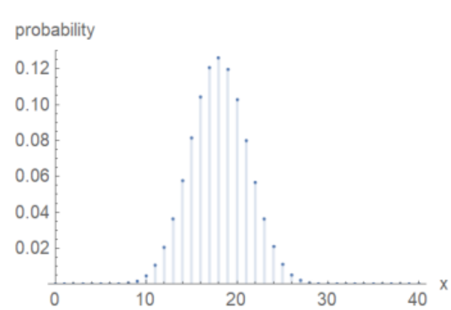
$X \sim$ Bernoulli with $p = 0.45$, or equivalently $X \sim \text{Bin}(1, 0.45)$.

$E(X) = 0.45$, $\sigma_X = \sqrt{(0.45)(0.55)} = 0.497$

(b) Let Y be the number of scam calls in the next 40 phone calls. What type of random variable is Y? Sketch the pmf of Y.

$Y \sim \text{Bin}(40, 0.45)$

pmf:



(c) What are the mean and standard deviation of Y?

$E(Y) = 40(0.45) = 18$

$\sigma_Y = \sqrt{40(0.45)(0.55)} = 3.14$

(d) Suppose that you lose 30 seconds of your time every time a scammer calls your phone. What are the expected value and standard deviation of the amount of time you will lose over the next 40 phone calls?

Let $Z = 30Y$ be the number of seconds you lose.

Then $E(Z) = 30\,E(Y) = 540$ seconds, and $\sigma_Z = 30\,\sigma_Y = 94$ seconds.

4. A coin that lands on heads with probability p is flipped ten times. Given that a total of 6 heads results, what is the conditional probability that the first three flips are heads, tails, heads (in that order)?

We will talk about this next time.

5. Among persons donating blood to a clinic, 85% have Rh^+ blood. Six people donate blood at the clinic on a particular day.

(a) Find the probability that at most three of the six have Rh^+ blood.

Next time...

(b) Find the probability that at most one of the six does not have Rh^+ blood.

(c) What is the probability that the number of Rh^+ donors lies within two standard deviations of the mean number?

(d) The clinic needs six Rh^+ donors on a certain day. How many people must donate blood to have the probability of obtaining blood from at least six Rh^+ donors over 0.95?

BONUS: Let $X \sim \text{Bin}(n, p)$. Show that $E(X) = np$.

$$E(X) = \sum_{x=1}^{n} x \binom{n}{x} p^{x} (1-p)^{n-x} = np \sum_{x=1}^{n} x \frac{(n-1)!}{x!\,(n-x)!} p^{x-1} (1-p)^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} (1-p)^{n-x}$$

let $j = x-1$

$$= np \sum_{j=0}^{n-1} \frac{(n-1)!}{j!\,(n-j-1)!} p^{j} (1-p)^{n-j-1}$$

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^{j} (1-p)^{n-j-1}$$

binomial theorem

$$= np \left(p + (1-p)\right)^{n-j-1}$$

$$= np\,(1)^{n-j-1} = np$$

BONUS: A system consists of n components, each of which will independently function with probability p. The system will operate effectively if at least one-half of its components function. For what values of p is a 5-component system more likely to operate than a 3-component system?

Let $X \sim \text{Bin}(5, p)$. The probability that a 5-component system functions effectively is $P(X \geq 3)$.

Similarly, let $Y \sim \text{Bin}(3, p)$. The probability that a 3-component system functions effectively is $P(Y \geq 2)$.

Thus, we want p such that:

$$P(X \geq 5) > P(Y \geq 2)$$

$$P(X=3) + P(X=4) + P(X=5) > P(Y=2) + P(Y=3)$$

$$10p^3(1-p)^2 + 5p^4(1-p) + p^5 > 3p^2(1-p) + p^3$$

Simplify to obtain:

$$3(p-1)^2(2p-1) > 0$$

$$p > \frac{1}{2}$$

A 5-component system is more likely than a 3-component system to operate effectively if $p > \frac{1}{2}$.