STA 291
Fall 2009

LECTURE 18
TUESDAY, 27 OCTOBER

Homework

- Graded online homework is due Saturday.

- Suggested problems from the textbook:

21.1 to 21.4,21.7 , 21.8, 21.10, and 21.11

Population Distribution vs. Probability Distribution

- If you select a subject randomly from the population, then the probability distribution for the value of the random variable X is the relative frequency (population, if you have it, but usually approximated by the sample version) of that value

Cumulative Distribution Function

Definition: The *cumulative distribution function*, or *CDF* is

$$F(x) = P(X \leq x).$$

Motivation: Some parts of the previous example would have been easier with this tool.

Properties:

1. For any value x, $0 \leq F(x) \leq 1$.
2. If $x_1 < x_2$, *then* $F(x_1) \leq F(x_2)$
3. $F(-\infty) = 0$ *and* $F(\infty) = 1$.

Example

Let X have the following probability distribution:

X	2	4	6	8	10
$P(x)$	.05	.20	.35	.30	.10

a.) Find $P(X \leq 6)$

b.) Graph the cumulative probability distribution of X

c.) Find $P(X > 6)$

Expected Value of a Random Variable

- The Expected Value, or mean, of a random variable, X, is
  Mean = E(X)= $\mu = \sum x_i P\left(X = x_i\right)$

- Back to our previous example—what's E(X)?

X	2	4	6	8	10
$P(x)$	.05	.20	.35	.30	.10

Useful formula

Suppose X is a random variable and a and b are constants. Then

$$E(a X + b) = aE(X) + b$$

Suppose X and Y are random variables and a, b, c are constants. Then

$$E(a X + b Y + c) = aE(X) + bE(Y) + c$$

Variance of a Random Variable

- Variance= Var(X) =

$$\sigma^2 = E\left[\left(X-\mu\right)^2\right] = \sum\left(x_i-\mu\right)^2 \cdot P\left(X=x_i\right)$$

- Back to our previous example—what's Var(X)?

X	**2**	**4**	**6**	**8**	**10**
***P*(*x*)**	.05	.20	.35	.30	.10

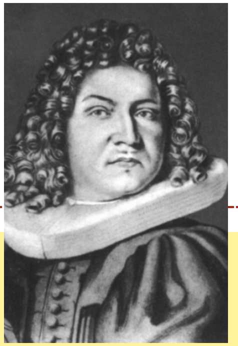
Useful formula

Suppose X is a random variable and a and b are constants then

$$Var(aX + b) = a^2 \, Var(X)$$

If X and Y are independent random variables then

$$Var(X + Y) = Var(X) + Var(Y)$$



Bernoulli Trial

- Suppose we have a single random experiment X with two outcomes: "success" and "failure."
- Typically, we denote "success" by the value 1 and "failure" by the value 0.
- It is also customary to label the corresponding probabilities as:

$$P(\text{success}) = P(1) = p \text{ and}$$

$$P(\text{failure}) = P(0) = 1 - p = q$$

- Note: $p + q = 1$

Binomial Distribution I

- Suppose we perform several Bernoulli experiments and they are all independent of each other.
- Let's say we do n of them. The value n is the **number of trials.**
- We will label these n Bernoulli random variables in this manner: $X_1, X_2, \ldots, X_n$
- As before, we will assume that the probability of success in a single trial is p, and that this probability of success doesn't change from trial to trial.

Binomial Distribution II

- Now, we will build a new random variable X using all of these Bernoulli random variables:

$$X = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

- What are the possible outcomes of X?
- What is X counting?
- How can we find $P(X = x)$?

Binomial Distribution III

- We need a quick way to count the number of ways in which *k* successes can occur in *n* trials.
- Here's the formula to find this value:

$$\binom{n}{k} = {}_nC_k = \frac{n!}{k!(n-k)!}, \text{ where } n! = n \cdot (n-1) \cdot \cdots \cdot 3 \cdot 2 \cdot 1 \text{ and } 0! \equiv 1$$

- Note: ${}_nC_k$ *is read as* "*n* choose *k*."

Binomial Distribution IV

- Now, we can write the formula for the binomial distribution:
- The probability of observing *x* successes in *n* independent trials is

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, \cdots, n$$

under the assumption that the probability of success in a single trial is *p*.

Using Binomial Probabilities

Note: Unlike generic random variables where we would have to be given the probability distribution or calculate it from a frequency distribution, here we can calculate it from a mathematical formula.

Helpful resources (besides your calculator):

- Excel:

Enter	Gives
=BINOMDIST(4,10,0.2,FALSE)	0.08808
=BINOMDIST(4,10,0.2,TRUE)	0.967207

- Table 1, pp. B-1 to B-5 in the back of your book

Binomial Probabilities

We are choosing a random sample of n = 7 Lexington residents—our random variable, C = number of Centerpointe supporters in our sample. Suppose, p = P (Centerpointe support) ≈ 0.3. Find the following probabilities:

a) $P (C = 2)$
b) $P (C < 2)$
c) $P (C \leq 2)$
d) $P (C \geq 2)$
e) $P (1 \leq C \leq 4)$

What is the *expected* number of Centerpointe supporters, μ_C?

Attendance Question #18

Write your name and section number on your index card.

Today's question: