

3.5–3.6 Expected values and variance

Prof. Tesler

Math 186
January 28, 2009

Expected Winnings in a Game

Setup

- A simple game:
 - 1 Pay \$1 to play once.
 - 2 Flip two fair coins.
 - 3 Win \$5 if HH, nothing otherwise.
- The **payoff** is $X = \begin{cases} \$5 & \text{with probability } 1/4; \\ \$0 & \text{with probability } 3/4. \end{cases}$
- The **net winnings** are
$$Y = X - 1 = \begin{cases} \$5 - \$1 = \$4 & \text{with probability } 1/4; \\ \$0 - \$1 = -\$1 & \text{with probability } 3/4. \end{cases}$$
- Playing the game once is called a **trial**.
- Playing the game n times is **an experiment with n trials**.

Expected Winnings in a Game

Average payoff

- If you play the game n times, the payoff will be \$5 about $n/4$ times and \$0 about $3n/4$ times, totalling

$$\$5 \cdot n/4 + \$0 \cdot 3n/4 = \$5n/4$$

- The **expected payoff** (long term average payoff) per game is obtained by dividing by n :

$$E(X) = \$5 \cdot 1/4 + \$0 \cdot 3/4 = \$1.25$$

- For the **expected winnings** (long term average winnings), subtract off the bet:

$$E(Y) = E(X - 1) = \$4 \cdot 1/4 - \$1 \cdot 3/4 = \$0.25$$

- That's good for you and bad for the house.
- A **fair game** has expected winnings = \$0.
A game **favors the player** if the expected winnings are positive.
A game **favors the house** if the expected winnings are negative.

Expected Value of a Random Variable

(Technical name for long term average)

- The **expected value** of a discrete random variable X is

$$E(X) = \sum_x x \cdot p_X(x)$$

- The **expected value** of a continuous random variable X is

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

- $E(X)$ is often called the **mean value of X** and is denoted μ (or μ_X if there is more than one random variable in the problem).
- μ doesn't have to be a value in the range of X . The previous example had range $X = \$0$ or $\$5$, and mean $\$1.25$.

Expected Value of Binomial Distribution

- Consider a biased coin with probability $p = 3/4$ for heads.
- Flip it 10 times and record the number of heads, x_1 .
Flip it another 10 times, get x_2 heads.
Repeat to get x_1, \dots, x_{1000} .
- **Estimate the average of x_1, \dots, x_{1000} : $10(3/4) = 7.5$**
(Later we'll show $E(X) = np$ for the binomial distribution.)
- **An estimate based on the pdf:**
About $1000p_X(k)$ of the x_i 's equal k for each $k = 0, \dots, 10$, so

$$\text{average of } x_i\text{'s} = \frac{\sum_{i=1}^{1000} x_i}{1000} \approx \frac{\sum_{k=0}^{10} k \cdot 1000 p_X(k)}{1000} = \sum_{k=0}^{10} k \cdot p_X(k)$$

which is the formula for $E(X)$ in this case.

Interpretation of the word “Expected”

- Although $E(X) = 7.5$, this is not a possible value for X .
- Expected value does *not* mean we anticipate observing that value.
- It means the long term average of many independent measurements of X will be approximately $E(X)$.

Function of a Random Variable

- Let X be the value of a roll of a biased die and $Z = (X - 3)^2$.

x	$p_X(x)$	$z = (x - 3)^2$	$p_Z(z)$
1	q_1	4	
2	q_2	1	
3	q_3	0	$p_Z(0) = q_3$
4	q_4	1	$p_Z(1) = q_2 + q_4$
5	q_5	4	$p_Z(4) = q_1 + q_5$
6	q_6	9	$p_Z(9) = q_6$

- pdf of X :** Each $q_i \geq 0$ and $q_1 + \cdots + q_6 = 1$.
- pdf of Z :** Each probability is also ≥ 0 , and the total sum is also 1.

Expected Value of a Function

- Let X be the value of a roll of a biased die and $Z = (X - 3)^2$.

x	$p_X(x)$	$z = (x - 3)^2$	$p_Z(z)$
1	q_1	4	
2	q_2	1	
3	q_3	0	$p_Z(0) = q_3$
4	q_4	1	$p_Z(1) = q_2 + q_4$
5	q_5	4	$p_Z(4) = q_1 + q_5$
6	q_6	9	$p_Z(9) = q_6$

- $E(Z)$, in terms of values of Z and the pdf of Z , is

$$E(Z) = \sum_z z \cdot p_Z(z) = 0(q_3) + 1(q_2 + q_4) + 4(q_1 + q_5) + 9(q_6)$$

- Regroup it in terms of X :

$$= 4q_1 + 1q_2 + 0q_3 + 1q_4 + 4q_5 + 9q_6 = \sum_{x=1}^6 (x - 3)^2 p_X(x)$$

Expected Value of a Function

- Let X be a discrete random variable, and $g(X)$ be a function, such as $(X - 3)^2$.
- The **expected value of $g(X)$** is

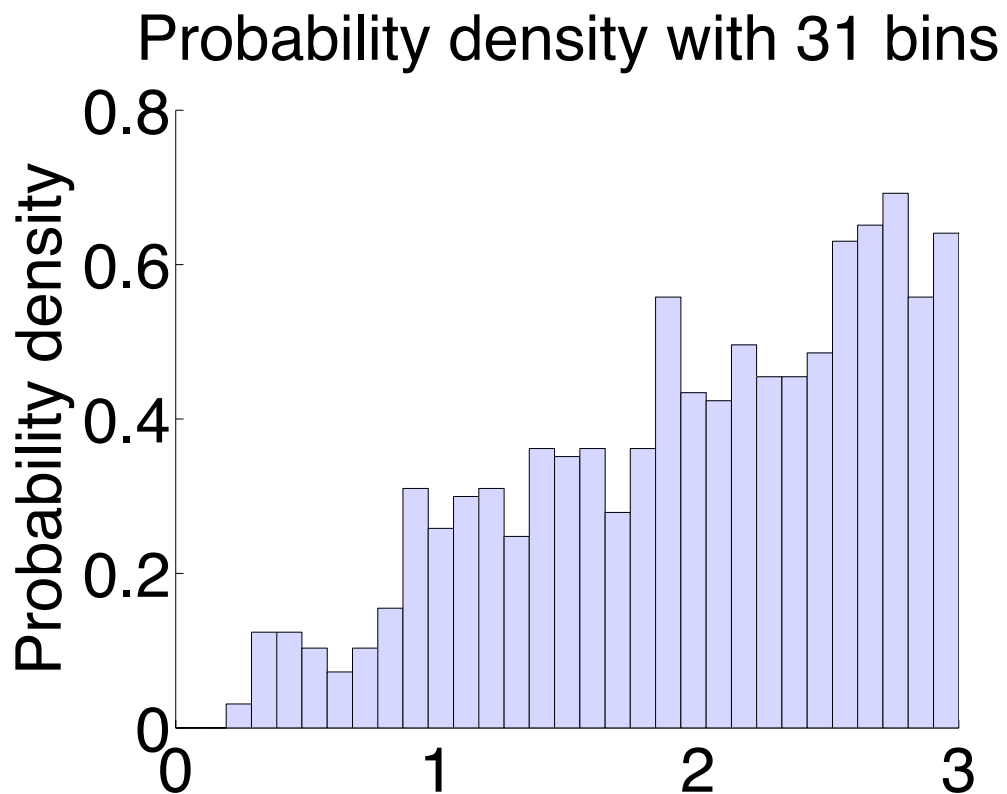
$$E(g(X)) = \sum_x g(x) p_X(x)$$

- For a continuous random variable,

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

- Note that if $Z = g(X)$ then $E(Z) = E(g(X))$.

Expected Value of a Continuous distribution



- Consider the dartboard of radius 3 example, with pdf

$$f_R(r) = \begin{cases} 2r/9 & \text{if } 0 \leq r \leq 3; \\ 0 & \text{otherwise.} \end{cases}$$

- Throw n darts and make a histogram with k bins.
- r_1, r_2, \dots are representative values of R in each bin.
- The bin width is $\Delta r = 3/k$, the height is $\approx f_R(r_i)$, and the area is $\approx f_R(r_i) \Delta r$.
- The approximate number of darts in bin i is $n f_R(r_i) \Delta r$.

Expected Value of a Continuous Distribution

- The estimated average radius is

$$\frac{\sum_i r_i \cdot n f_R(r_i) \Delta r}{n} = \sum_i r_i \cdot f_R(r_i) \Delta r$$

- As $n, k \rightarrow \infty$, the histogram smoothes out and this becomes

$$\int_0^3 r \cdot f_R(r) dr$$

Mean of a continuous distribution

- Consider the dartboard of radius 3 example, with pdf

$$f_R(r) = \begin{cases} 2r/9 & \text{if } 0 \leq r \leq 3; \\ 0 & \text{otherwise.} \end{cases}$$

- The “average radius” (technically the **mean radius** or **expected value of R**) is

$$\begin{aligned} \mu = E(R) &= \int_{-\infty}^{\infty} r \cdot f_R(r) \, dr = \int_0^3 r \cdot \frac{2r}{9} \, dr = \int_0^3 \frac{2r^2}{9} \, dr \\ &= \left. \frac{2r^3}{27} \right|_0^3 = \frac{2(3^3 - 0^3)}{27} = 2 \end{aligned}$$

Expected Values — Properties

- The gambling slide earlier had $E(X - 1) = E(X) - 1$.

Theorem

$$E(aX + b) = aE(X) + b \quad \text{where } a, b \text{ are constants.}$$

Proof (discrete case).

$$\begin{aligned} E(aX + b) &= \sum_x (ax + b) \cdot p_X(x) \\ &= a \sum_x x \cdot p_X(x) + b \sum_x p_X(x) \\ &= a \cdot E(X) + b \cdot 1 = aE(X) + b \end{aligned}$$



Expected Values — Properties

- The gambling slide earlier had $E(X - 1) = E(X) - 1$.

Theorem

$$E(aX + b) = aE(X) + b \quad \text{where } a, b \text{ are constants.}$$

Proof (continuous case).

$$\begin{aligned} E(aX + b) &= \int_{-\infty}^{\infty} (ax + b) \cdot f_X(x) \, dx \\ &= a \int_{-\infty}^{\infty} x \cdot f_X(x) \, dx + b \int_{-\infty}^{\infty} f_X(x) \, dx \\ &= a \cdot E(X) + b \cdot 1 = aE(X) + b \end{aligned}$$



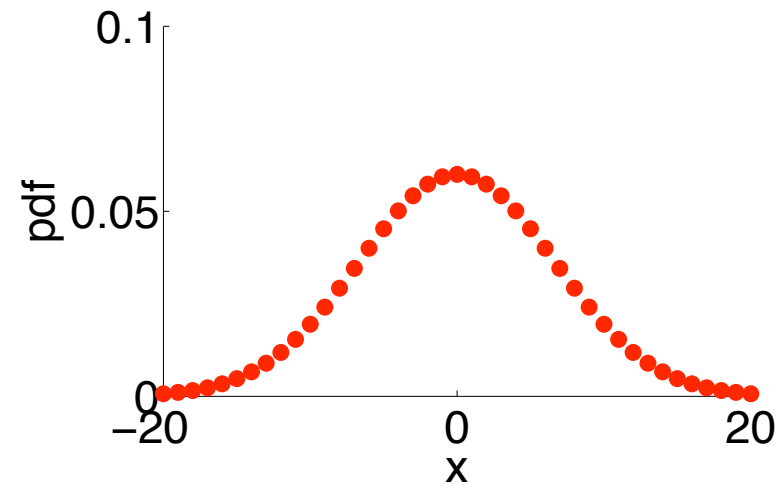
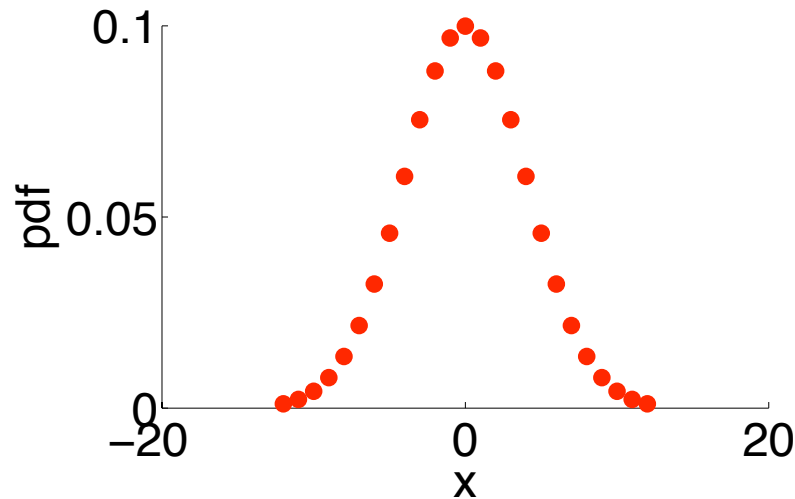
Expected Values — Properties

These properties hold for both discrete and continuous random variables:

- $E(aX + b) = aE(X) + b$ for any constants a, b .
- $E(aX) = aE(X)$
- $E(b) = b$
- $E(g(X) + h(X)) = E(g(X)) + E(h(X))$

Variance

- These distributions both have mean=0, but the right one is more spread out.



- The **variance** of X measures the square of the spread from the mean:

$$\sigma^2 = \text{Var}(X) = E((X - \mu)^2)$$

- The **standard deviation** of X is $\sigma = \sqrt{\text{Var}(X)}$ and measures how wide the curve is.

Variance formula $\sigma^2 = E((X - \mu)^2)$

- Consider the dartboard of radius 3 example, with pdf

$$f_R(r) = \begin{cases} 2r/9 & \text{if } 0 \leq r \leq 3; \\ 0 & \text{otherwise.} \end{cases}$$

- $\mu = 2$ from earlier slide.

- $$\begin{aligned} \sigma^2 = \text{Var}(R) &= E((R - \mu)^2) = E((R - 2)^2) \\ &= \int_{-\infty}^{\infty} (r - 2)^2 f_R(r) dr = \int_0^3 \frac{(r - 2)^2 \cdot 2r}{9} dr \\ &= \int_0^3 \frac{2r^3 - 8r^2 + 8r}{9} dr = \left(\frac{r^4}{18} - \frac{8r^3}{27} + \frac{4r^2}{9} \right) \Big|_0^3 \\ &= \left(\frac{3^4}{18} - \frac{8(3^3)}{27} + \frac{4(3^2)}{9} \right) - 0 = \frac{1}{2} \end{aligned}$$

- **Variance:** $\sigma^2 = \frac{1}{2}$ **Standard deviation:** $\sigma = \sqrt{1/2}$

Variance — Second formula

There are two equivalent formulas to compute variance.

In any problem, choose the easier one:

$$\begin{aligned}\sigma^2 = \text{Var}(X) &= E((X - \mu)^2) \quad (\text{Definition}) \\ &= E(X^2) - \mu^2 \quad (\text{Sometimes easier to compute})\end{aligned}$$

Proof.

$$\begin{aligned}\text{Var}(X) &= E((X - \mu)^2) \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2\end{aligned}$$



Variance formula $\sigma^2 = E(R^2) - \mu^2$

- Consider the dartboard of radius 3 example, with pdf

$$f_R(r) = \begin{cases} 2r/9 & \text{if } 0 \leq r \leq 3; \\ 0 & \text{otherwise.} \end{cases}$$

- $\mu = E(R) = 2$

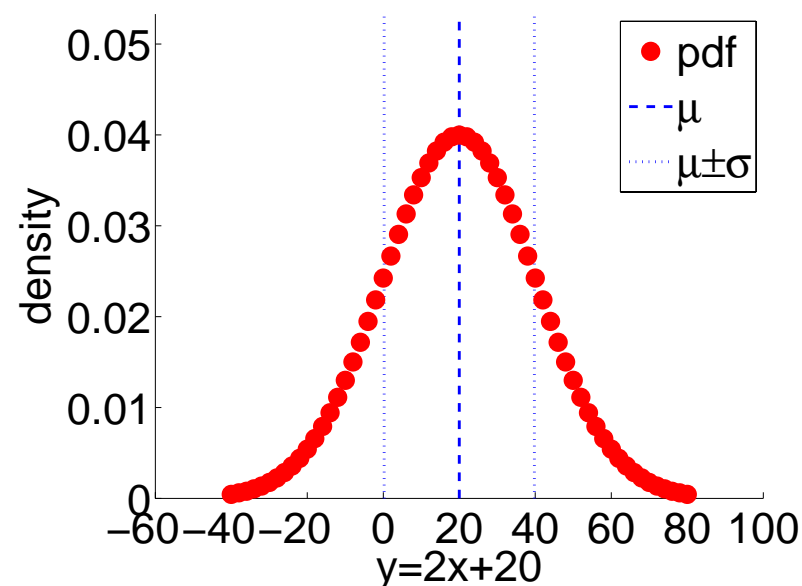
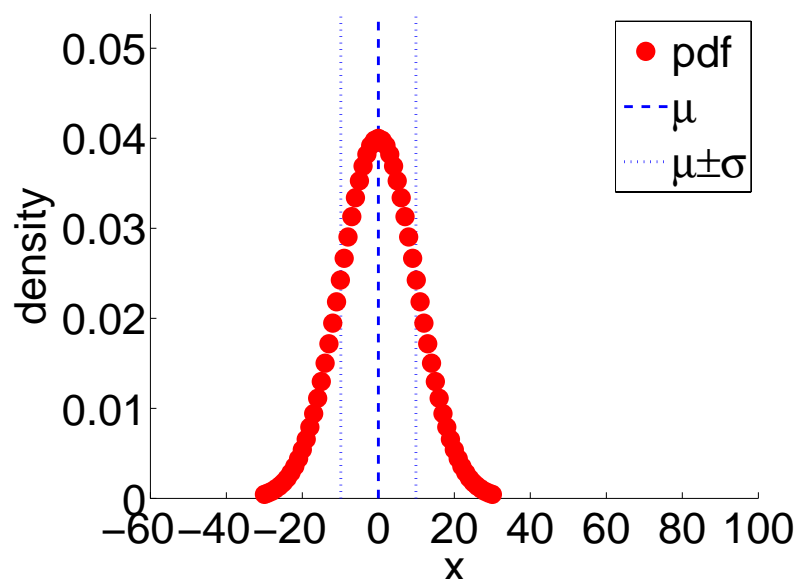
- $E(R^2) = \int_0^3 r^2 \cdot \frac{2r}{9} dr = \int_0^3 \frac{2r^3}{9} dr = \frac{2r^4}{36} \Big|_0^3 = \frac{2(81 - 0)}{36} = 9/2$

- **Variance:** $\sigma^2 = E(R^2) - \mu^2 = \frac{9}{2} - 2^2 = \frac{1}{2}$

Standard deviation: $\sigma = \sqrt{1/2}$

Variance — Properties

- $\text{Var}(aX + b) = a^2 \text{Var}(X)$



- Adding b shifts the curve without changing the width, so b disappears on the right side of the formula.

- Multiplying by a dilates the width a factor of a , so variance goes up a factor a^2 .

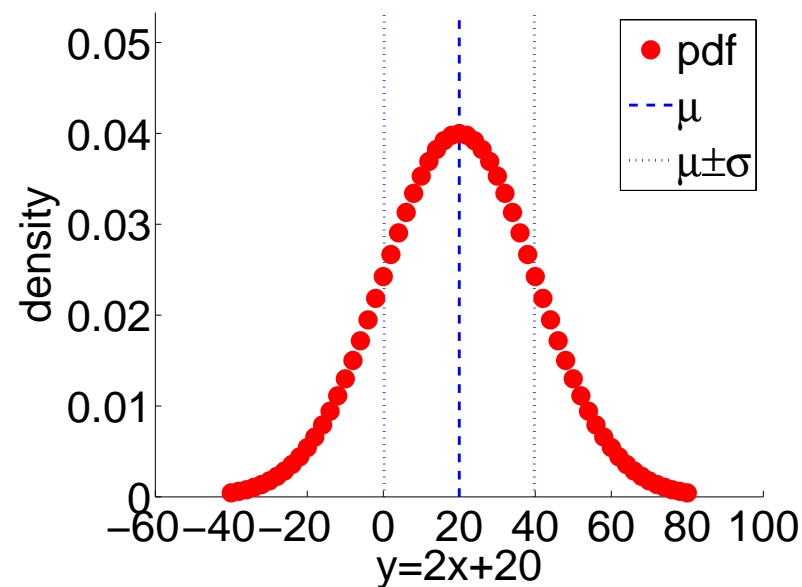
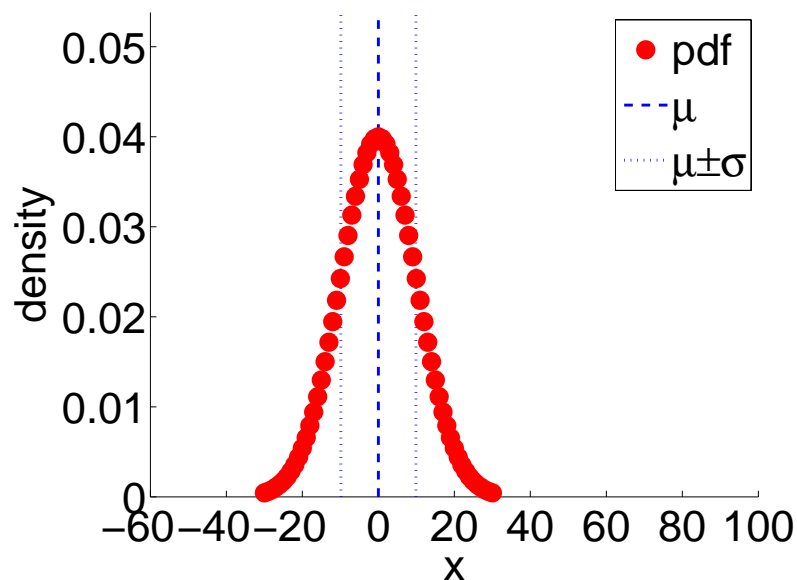
- For $Y = aX + b$, we have $\sigma_Y = |a| \sigma_X$.

- **Example:** Convert measurements in $^{\circ}C$ to $^{\circ}F$:

$$F = (9/5)C + 32 \quad \mu_F = (9/5)\mu_C + 32 \quad \sigma_F = (9/5)\sigma_C$$

Variance — Properties

- $\text{Var}(aX + b) = a^2 \text{Var}(X)$



Proof of $\text{Var}(aX + b) = a^2 \text{Var}(X)$.

$$\begin{aligned} E((aX + b)^2) &= E(a^2X^2 + 2abX + b^2) = a^2E(X^2) + 2abE(X) + b^2 \\ (E(aX + b))^2 &= (aE(X) + b)^2 = a^2(E(X))^2 + 2abE(X) + b^2 \\ \text{Var}(aX + b) &= \text{difference} = a^2(E(X^2) - (E(X))^2) \\ &= a^2 \text{Var}(X) \quad \square \end{aligned}$$

Mean and Variance of the Binomial Distribution

- For the binomial distribution,

Mean: $\mu = np$

Standard deviation:

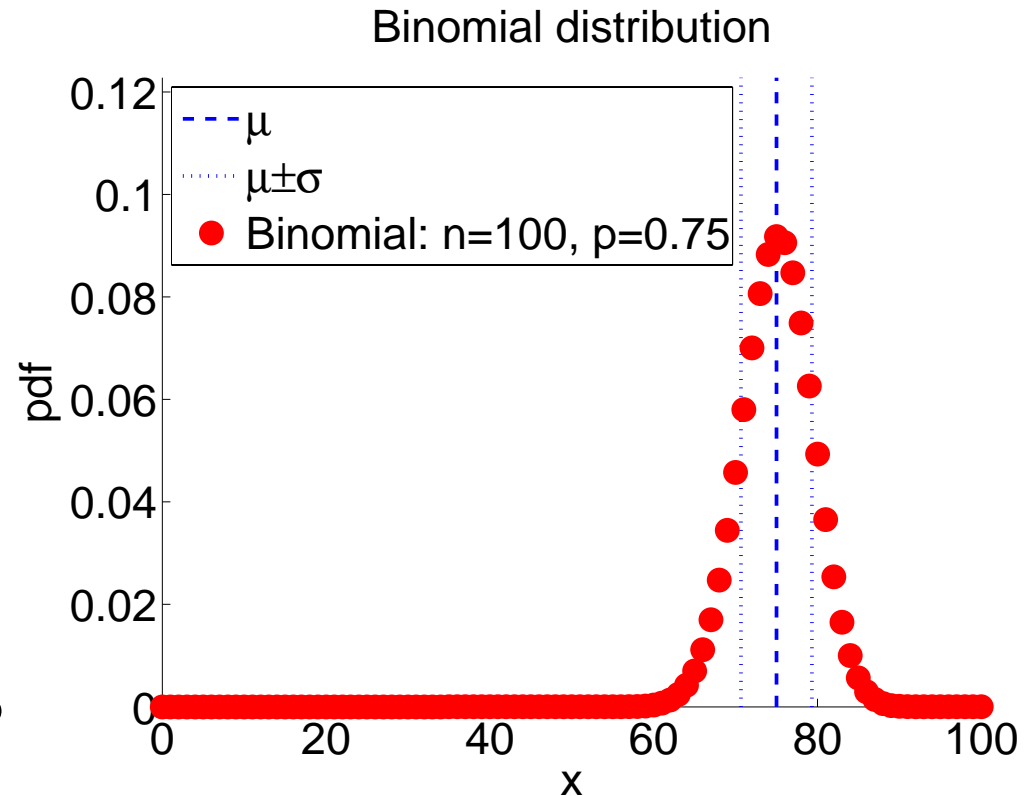
$$\sigma = \sqrt{np(1-p)}$$

- At $n = 100$ and $p = 3/4$:

$$\mu = 100(3/4) = 75$$

$$\sigma = \sqrt{100(3/4)(1/4)} \approx 4.33$$

- Approximately 68% of the probability is for X between $\mu \pm \sigma$.
Approximately 95% of the probability is for X between $\mu \pm 2\sigma$.
More on that in Chapter 4.



Mean of the Binomial Distribution

Proof that $\mu = np$ for binomial distribution.

$$\begin{aligned} E(X) &= \sum_k k \cdot p_X(k) \\ &= \sum_{k=0}^n k \cdot \binom{n}{k} p^k q^{n-k} \end{aligned}$$

Calculus Trick: $(p + q)^n = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$

Differentiate: $\frac{\partial}{\partial p} (p + q)^n = \sum_{k=0}^n k \binom{n}{k} p^{k-1} q^{n-k}$

Times p : $p \frac{\partial}{\partial p} (p + q)^n = \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k} = E(X)$

Evaluate it: $p \frac{\partial}{\partial p} (p + q)^n = p \cdot n(p + q)^{n-1}$
 $= p \cdot n \cdot 1^{n-1} = np$ since $p + q = 1$.

So $E(X) = np$. □

We'll do $\sigma = \sqrt{np(1-p)}$ later.