

We will use the notation $X \sim \mathcal{B}(n, p)$ to say X is binomially distributed with n trials and chance of success p (on each trial).

3.25: (a): Yes. Each trial is independent and has the same probability of success (and each trial will either be a success or failure).

(b): Let $X \sim \mathcal{B}(10, 0.697)$. We are asked to find the probability $P(X = 6)$.

$$\begin{aligned} P(X = 6) &= \binom{10}{6} (0.697)^6 (1 - 0.697)^4 \\ &= 210 \times 0.697^6 \times 0.303^4 \\ &= \boxed{0.203} \end{aligned}$$

(c): This is the same thing as part (b), just worded differently. Each child either did or didn't. If 6 did, then 4 didn't.

$$\boxed{0.203}$$

(d): Let $X \sim \mathcal{B}(n = 5, p = 0.697)$. We can calculate the entire probability mass function.

k	$P(X = k)$ before simplification	$P(X = k)$
0	$\binom{5}{0} 0.697^0 \times 0.303^5$	0.00255395
1	$\binom{5}{1} 0.697^1 \times 0.303^4$	0.02937469
2	$\binom{5}{2} 0.697^2 \times 0.303^3$	0.13514297
3	$\binom{5}{3} 0.697^3 \times 0.303^2$	0.31087342
4	$\binom{5}{4} 0.697^4 \times 0.303^1$	0.35755573
5	$\binom{5}{5} 0.697^5 \times 0.303^0$	0.16449924

We are asked to determine the probability $P(X \leq 2)$.

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= 0.00255 + 0.02937 + 0.13514 \\ &\approx \boxed{0.167} \end{aligned}$$

Also, if you are interested... R will evaluate this with a single command:

```
pbinom(2, 5, 0.697)
```

Also, we can use `dbinom(k,n,p)` to get the simple probabilities.

```
> dbinom(seq(0,5), 5, 0.697)
[1] 0.00255395 0.02937469 0.13514297 0.31087342 0.35755573 0.16449924
> pbinom(2, 5, 0.697)
[1] 0.1670716
```

(e): (Same distribution as (d).) It is easiest to recognize the complement of “at least 1” is “none”.

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - 0.00255 \\ &= \boxed{0.99745} \end{aligned}$$

3.26: (a): Yes. Each trial will be “success” or “fail”, and each trial has the same probability of “success”.

(b): Let $X \sim \mathcal{B}(100, 0.90)$. We are asked to calculate the probability $P(X = 97)$.

$$\begin{aligned} P(X = 97) &= \binom{100}{97} (0.9)^{97} (0.1)^3 \\ &= 161700 \times 0.9^{97} \times 0.1^3 \\ &= \boxed{0.005891602} \end{aligned}$$

We could also use R.

```
> dbinom(97, 100, 0.9)
[1] 0.005891602
```

(c): This is the same question as (b).

$$\boxed{0.00589}$$

(d): Let $Y \sim \mathcal{B}(10, 0.9)$, where Y is the number of people who had chickenpox from a random sample of 10 Americans. We recognize “at least one” as the complement of “none”.

$$\begin{aligned} P(Y = 0) &= \binom{10}{0} (0.9)^0 (0.1)^{10} \\ &= 0.000000001 \end{aligned}$$

We use the complement rule.

$$\begin{aligned} P(Y \geq 1) &= 1 - P(Y = 0) \\ &= 1 - 0.000000001 \\ &= 0.999999999 \\ &\approx \boxed{1} \end{aligned}$$

(e): We continue to let $Y \sim \mathcal{B}(10, 0.9)$. We rephrase the question in terms of number of people who *had* chickenpox. We want to find the probability that *at least 7* Americans had chickenpox.

$$\begin{aligned} P(Y \geq 7) &= P(Y = 7) + P(Y = 8) + P(Y = 9) + P(Y = 10) \\ &= \binom{10}{7} (0.9)^7 (0.1)^3 + \binom{10}{8} (0.9)^8 (0.1)^2 + \binom{10}{9} (0.9)^9 (0.1)^1 + \binom{10}{10} (0.9)^{10} (0.1)^0 \\ &\approx \boxed{0.9872} \end{aligned}$$

3.27: Let $X \sim \mathcal{B}(50, 0.70)$. As a reminder, that means $n = 50$ and $p = 0.7$.

(a): We can calculate the mean and standard deviation of X using the formulas for binomial distributions.

$$\mu = np = 50 \times 0.70 = 35$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(50)(0.7)(0.3)} = 3.24$$

(b): Well, 45 would have a pretty high z score. Remember, we can often approximate a binomial distribution with a normal distribution.

$$z = \frac{45 - 35}{3.24} = 3.09$$

Yeah. If $|z| > 2$, we consider it unusual, as it happens less than 5% of the time due to chance. If $|z| > 3$, we consider it rare, as it happens less than 0.3% of the time due to chance.

(c): Hmm... we can do this a few ways. The first is to use R.

```
> 1 - pbinom(44, 50, 0.7)
[1] 0.0007228617
> sum(dbinom(seq(45, 50), 50, 0.7))
[1] 0.0007228617
```

We can do this by hand...

k	$P(X = k)$ before simplification	$P(X = k)$
\vdots	\vdots	\vdots
45	$\binom{50}{45}(0.7)^{45}(0.3)^5$	0.000551
46	$\binom{50}{46}(0.7)^{46}(0.3)^4$	0.000140
47	$\binom{50}{47}(0.7)^{47}(0.3)^3$	0.000028
48	$\binom{50}{48}(0.7)^{48}(0.3)^2$	0.000004
49	$\binom{50}{49}(0.7)^{49}(0.3)^1$	0.000000
50	$\binom{50}{50}(0.7)^{50}(0.3)^0$	0.000000

We sum the probabilities to get $P(X \geq 45) = 0.00072$.

As a third technique, we can use the **normal approximation** with **continuity correction**.

Let $Y \sim \mathcal{N}(35, 3.24)$.

$$\begin{aligned}
 P(X \geq 45) &\approx P(Y \geq 44.5) \\
 &= 1 - \Phi\left(\frac{44.5 - 35}{3.24}\right) \\
 &= 0.00168
 \end{aligned}$$

But, notice the normal approximation did not work out great... it was off by a factor of more than 2. **The normal approximation does not work well in the tails.** However, all three techniques agree that the probability is quite low.

3.28: (a): Let $X \sim \mathcal{B}(120, 0.9)$. To determine the mean and standard deviation, we use the formulas for binomial distributions.

$$\begin{aligned}\mu &= np = (120)(0.9) = 108 \\ \sigma &= \sqrt{np(1-p)} = \sqrt{(120)(0.9)(0.1)} = 3.286\end{aligned}$$

We expect about 108 with a standard deviation of 3.28 people to have had chickenpox.

(b): 105 is within 1 standard deviation of 108.

$$z = \frac{105 - 108}{3.28} = -0.914$$

So, this does not seem unusual.

(c): We are asked to calculate $P(X \leq 105)$. This is easy with a computer.

```
> pbinom(105, 120, 0.9)
[1] 0.2181634
```

To do it by hand, we will use a normal approximation with the continuity correction. Let $Y \sim \mathcal{N}(108, 3.286)$.

$$\begin{aligned}P(X \leq 105) &\approx P(Y \leq 105.5) \\ &= \Phi\left(\frac{105.5 - 108}{3.286}\right) \\ &= 0.2233872\end{aligned}$$

So, either way we get about 22% chance, which agrees with us not being surprised in part (b).

3.29: Let $X \sim \mathcal{B}(2500, 0.7)$, where X represents the number of students who accept. We will definitely use a normal approximation here. In fact, we are dealing with such a large n we don't even need to make a continuity correction.

We first calculate the mean and standard deviation using the formulas for binomial distributions.

$$\begin{aligned}\mu &= np = 2500 \times 0.7 = 1750 \\ \sigma &= \sqrt{np(1-p)} = \sqrt{2500(0.7)(0.3)} = 22.9\end{aligned}$$

Now, let $Y \sim \mathcal{N}(1750, 22.9)$. The normal approximation (with continuity correction) can be used.

$$\begin{aligned}P(X > 1786) &\approx P(Y > 1786.5) \\ &= 1 - \Phi\left(\frac{1786.5 - 1750}{22.9}\right) \\ &= 0.056\end{aligned}$$

Let's use R to calculate it more exactly.

```
> 1-pbinom(1786, 2500, 0.7)
[1] 0.05506358
```

Wow, in this case the normal approximation did a great job!

3.30: Let $X \sim \mathcal{B}(15000, 0.09)$. Remember, that means $n = 15000$ and $p = 0.09$. We want to calculate the probability $P(X \geq 1500)$. Using a computer, we can do that easily.

```
> 1-pbinom(1499, 15000, 0.09)
[1] 1.326331e-05
```

Let's also use a normal approximation. We first calculate the mean and standard deviation.

$$\mu = np = (15000)(0.09) = 1350$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(15000)(0.09)(0.91)} = 35.04996$$

Now we define a continuous random variable $Y \sim \mathcal{N}(1350, 35)$, and we use the normal approximation.

$$\begin{aligned} P(X \geq 1500) &\approx P(Y > 1499.5) \\ &= 1 - \Phi\left(\frac{1499.5 - 1350}{35}\right) \\ &= 0.0000097 \end{aligned}$$

We don't expect the normal approximation to work in the tails. But we know the chance is low!