Michael is 190cm tall. In his high school, heights are normally distributed with a mean of 165 cm and a standard deviation of 20 cm. What is the probability that Michael's best friend is shorter than he is?

What's the difference?

At the height of The Beatles' popularity, it was estimated that their music was played on every popular music radio station 40 percent of the time. What is the probability that if you tuned through 1,000 such stations at any given moment at least 200 of the stations would be playing a Beatles song?

Reason 2:

Using the Normal to Approximate to the Binomial

Consider the following:

Quality Control at Widgets Inc. claims that 4% of their widgets are defective. What is the probability that 30 or less of a random sample of 875 widgets are defective?

Solving this as a binomial will require about 30 calculations.

$P(X<30) = P(X=0) + P(X=1) + ... + P(X=29)$

where

$P(X=0) = {}_{875}C_{0}(0.04)^{0}(0.96)^{875}$

$P(X=1) = {}_{875}C_{1}(0.04)^{1}(0.96)^{874}$

$P(X=2) = {}_{875}C_{2}(0.04)^{2}(0.96)^{873}$

<and so on until...>

$P(X=29) = {}_{875}C_{29}(0.04)^{29}(0.96)^{846}$

Time's wasting - get to it!

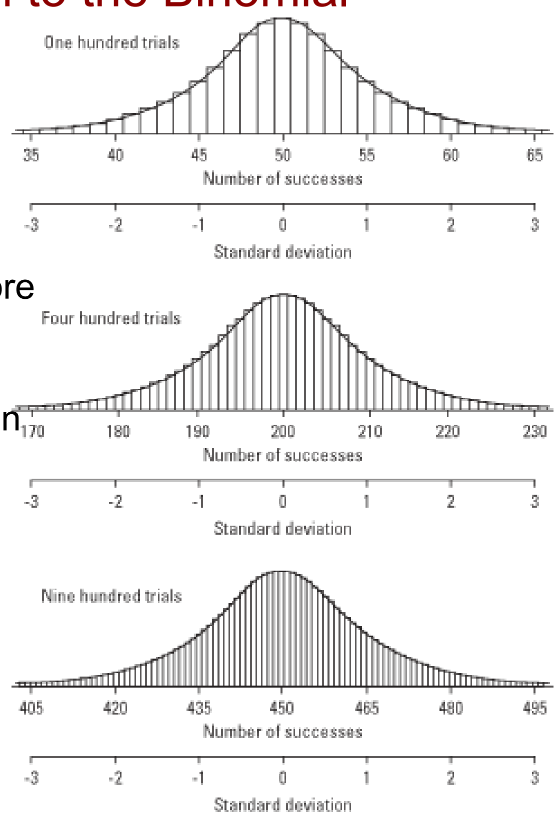
The Normal Approximation to the Binomial

If we knew the mean and the standard deviation, we could use z-scores and the normal curve.

There are conditions, though. Before using the Normal Approximation, calculate n×P and n×Q. If both of the answers are greater than 5, then the approximation can be used safely.

The next step is to calculate the approximate mean and standard deviation:

$\mu = n \times P$ and $\sigma = \sqrt{(nPQ)}$



Considering our first example from this perspective...

$$n \times p = 875 \times 0.04 \qquad n \times q = 875 \times 0.96$$
$$= 35 \qquad = 840$$
$$> 5 \qquad > 5$$

Since both of these values are greater than 5, the normal approximation is valid.

$$\mu = n \times p$$
$$= 875 \times 0.04$$
$$= 35$$

$$\sigma = \sqrt{n \times p \times q}$$
$$= \sqrt{875 \times 0.04 \times 0.96}$$
$$= \sqrt{33.6}$$
$$= 5.79$$

We can now calculate the relevant z-score using these values.

Quality Control insists that no more than 30 will be defective. This provides us with an x value.

$$z = \frac{x - \mu}{\sigma}$$
$$= \frac{30 - 35}{5.79}$$
$$= \frac{-5}{5.79}$$
$$= -0.86$$

From the z-score table, this corresponds to a probability of 19.49%.

x	n-x	875 C x	p^x	q^(n-x)	result
0	875	1	1	3.07135E-16	3.07135E-16
1	874	875	0.04	3.19932E-16	1.11976E-14
2	873	382375	0.0016	3.33262E-16	2.0389E-13
3	872	111271125	0.000064	3.47148E-16	2.47217E-12
4	871	2.4257E+10	0.00000256	3.61613E-16	2.24555E-11
5	870	4.2256E+12	1.024E-07	3.7668E-16	1.6299E-10
6	869	6.1271E+14	4.096E-09	3.92375E-16	9.84729E-10
7	868	7.6064E+16	1.6384E-10	4.08724E-16	5.09363E-09
8	867	8.2529E+18	6.5536E-12	4.25754E-16	2.30274E-08
9	866	7.9503E+20	2.6214E-13	4.43494E-16	9.24296E-08
10	865	6.885E+22	1.0486E-14	4.61973E-16	3.33517E-07
11	864	5.4141E+24	4.1943E-16	4.81222E-16	1.09277E-06
12	863	3.8981E+26	1.6777E-17	5.01273E-16	3.27832E-06
13	862	2.5878E+28	6.7109E-19	5.22159E-16	9.06791E-06
14	861	1.5933E+30	2.6844E-20	5.43916E-16	2.32635E-05
15	860	9.1457E+31	1.0737E-21	5.66579E-16	5.56386E-05
16	859	4.9158E+33	4.295E-23	5.90186E-16	0.000124607
17	858	2.4839E+35	1.718E-24	6.14777E-16	0.000262347
18	857	1.184E+37	6.8719E-26	6.40393E-16	0.00052105
19	856	5.3405E+38	2.7488E-27	6.67076E-16	0.000979255
20	855	2.2857E+40	1.0995E-28	6.94871E-16	0.001746337
21	854	9.3062E+41	4.398E-30	7.23824E-16	0.002962536
22	853	3.6125E+43	1.7592E-31	7.53983E-16	0.004791678
23	852	1.3398E+45	7.0369E-33	7.85399E-16	0.007404532
24	851	4.7562E+46	2.8147E-34	8.18124E-16	0.010952537
25	850	1.619E+48	1.1259E-35	8.52213E-16	0.015534348
26	849	5.2929E+49	4.5036E-37	8.87722E-16	0.02116057
27	848	1.6643E+51	1.8014E-38	9.2471E-16	0.027724265
28	847	5.0405E+52	7.2058E-40	9.6324E-16	0.034985382
29	846	1.4722E+54	2.8823E-41	1.00337E-15	0.042575602
					0.171817843

This screenshot from Excel shows the exact answer, according to the Binomial Distribution.

Practice Questions

In the exercises below, use the normal approximation to the binomial distribution, unless otherwise indicated. In each question, test that the normal approximation is valid by calculating np and nq – both must be greater than 5.

1. It is estimated that 62% of television viewers "channel surf" during commercials. A market-research firm surveyed 1500 television viewers. What is the probability that at least 950 of them were channel surfing?

2. Salespeople sometimes advertise their products by telephoning strangers. Only about 1.5% of these "cold calls" result in a sale. Toni makes cold calls 8h per day for 5 days. The average time for a cold call is 90s. What is the probability that Toni gets at least 30 new customers for the week?

3. A magazine reported that 18% of car drivers use a cellular phone while driving. In a survey of 200 drivers, what is the probability that exactly 40 of them will use a cellular phone while driving? Compare the results of using the binomial distribution and the normal approximation.

4. A recent survey of a gas station's customers showed that 68% paid with credit cards, 29% used debit cards, and only 3% paid with cash. During her eight hour shift as cashier at this gas station, Serena had a total of 223 customers.
   a) What is the probability that
      i) at least 142 customers used a credit card?
      ii) fewer than 220 customers paid with credit or debit cards?
   b) What is the expected number of customers who paid Serena with cash?

5. Calculate the probability that 200 rolls of two dice will include
   a) more than 30 sums of 5
   b) between 30 and 40, inclusive, sums of 5

6. On some busy streets, diamond lanes are reserved for taxis, buses and cars with three or more passengers. It is estimated that 20% of cars traveling in a certain diamond lane have fewer than three passengers. Sixty cars are selected at random.
   a) Use the normal approximation to find the probability that
      i) fewer than 10 cars have fewer than three passengers
      ii) at least 15 cars have fewer than three passengers
   b) Compare these results with those found using the binomial distribution.
   c) How would the results compare if 600 cars were selected?

7. IQ scores of people around the world are normally distributed, with a mean of 100 and a standard deviation of 15. A genius is someone with an IQ greater than or equal to 140. What percent of the population is considered genius?

8. A drug has a 70% success rate. What is the probability that 80 or more people out of 100 will be cured by the drug?

9. Mr. Median is a very precise teacher. Each class he has to have a class mean of 71 and a standard deviation of 11. What would the quartiles of Mr. Median's class be?

10. A snake farm advertises that 25% of their snakes are longer than 1.5 m and 10% of them are longer than 2 m. What is the mean length of snakes at this farm? What is the standard deviation?

Answer Clues

1) ~14.4%

2) ~11%

3) ~5.6% vs. ~5.4%

4ai) ~91.6%
aii) 93%
b) 6 or 7

5a) ~3.9%
b) ~3.9%

6ai) ~25.9%
aii) ~16.7%
b) 21.3% and 20.7%
c) normal: nearly 0% and 100%. Binomial: same

7) 3.8%

8) Normal: 1.5%. Binomial: 1.6%

9) approx. 63.6%, 71%, 78.4%, 100%

10) $\sigma = 0.826m$, $\mu = 0.94m$

Practice Assignment