Investigating a Probability Distribution

Kenn Pendelton

CALCULATORS: Casio: *fx-9750G Plus* • Casio: *CFX-9850G Series*

Teaching Notes/Lesson Plan

Setting the background

The hospital in City A averages 50 births per day. The hospital in Town B averages 20 births per day. Assume that the probabilities of boys and girls being born at both hospitals are equal and constant, and that births are independent. Based on the average births per day, the probability that 60% or more of the newborns on a given day will be girls is:

- Greater in City A than Town B
- Greater in Town B than City A
- Equal for City A and Town B

It is hoped that students will recognize that this could be a binomial distribution based on the two outcomes. Some suggestion in that direction may be necessary. Once that decision is reached, the conditions for a binomial distribution must be checked.

- two mutually exclusive and exhaustive outcomes (boy and girl)
- probability of success (girl births) remains constant (stated)
- trials are independent (stated)
- fixed number of trials (50 for City A, 20 for Town B)

Three approaches to the solution can be presented. All approaches will use counts, but it is perfectly acceptable to approach the question from the perspective of probabilities.

1. use the mean and standard deviation to compare relative probabilities of 60% or more girl births

2. compare the cumulative binomial probabilities of 60% or more girl births

3. approximate the binomial distribution with a normal distribution (after checking whether it is appropriate to do so) and compare the probabilities for 60% or more girl births

1. Relative probabilities

The mean of a binomial distribution (count) is np, and the standard deviation is $\sqrt{(npq)}$.

City A: $\mu = (50)(0.5) = 25$; $\sigma = \sqrt{[(50)(0.5)(0.5)]} \approx 3.54$; 60% of the births is 30.

Town B: $\mu = (20)(0.5) = 10$; $\sigma = \sqrt{[(20)(0.5)(0.5)]} \approx 2.24$; 60% of the births is 12.

For City A, 60% of the births (30) is MORE than one standard deviation (3.54) from the mean (25). For Town B, 60% of the births (12) is LESS than one standard deviation (2.24) from the mean (10).

CALCULATORS: Casio: *fx-9750G Plus* • Casio: *CFX-9850G Series*

Teaching Notes/Lesson Plan

The measure that is less extreme in terms of standard deviations is more likely to occur. Thus, 60% or more girl births are more likely to occur in Town B (where that is less than one standard deviation from the mean) than in City A (where that is more than one standard deviation from the mean).

2. Cumulative binomial probabilities

The cumulative binomial probability function calculates the cumulative probability from zero successes to a stated maximum, which is the left side of the probability distribution. To calculate the right side of the distribution, calculate its complement (girl births < 60%) and subtract from 1.

City A:
MENU; **STAT**; **DIST** (**F5**); **BINM** (**F5**); **Bcd** (**F2**); data from both lists and variables can be accommodated; select **Var** (***F2***); cursor down to ***x***, the maximum value; 29 (60% is 30; the greatest number of births less that 60% would be 29); **EXE**; Numtrial (total number of births) is 50; **EXE**; **p** (probability of success, a girl birth) is 0.5; **EXE**; **EXE** or **CALC** (**F1**). The cumulative probability of from 0 to 29 girl births at the hospital in City A is 0.89868. Therefore, the probability of 60% or more girl births is: 1 – 0.89868 or 0.10132.

Following the same procedure for Town B, the probability of 60% or more girl births is 0.25173. The probability is higher for Town B.

3. Approximating the binomial distribution with a normal distribution

Before using this method, students should determine whether it would be appropriate to do so. The typical approach is to determine whether both np and $nq \geq 10$.

City A: $np = nq = (50)(0.5) = 25$; this is acceptable.

Town B: $np = nq = (20)(0.5) = 10$; this is acceptable.

NOTE: The above approach is a generalization for the condition that $\mu \pm 3\sigma$ is wholly contained in the interval $[0, n]$. When $p = 0.5$, the restriction that np (and nq) ≥ 10 is very conservative.

CALCULATORS: Casio: *fx-9750G Plus* • Casio: *CFX-9850G Series*

Teaching Notes/Lesson Plan

The binomial distribution is a discrete distribution, but the normal is continuous. When approximating, the best value can be obtained by "filling in the space between the discrete values" by adding 0.5 to and subtracting 0.5 from each value.

City A: for the normal distribution, 30 (which is 60% of the births) is converted to the range of values 29.5 through 30.5. Greater than or equal to 30 for the binomial distribution is represented by greater than or equal to 29.5 (the least value for 30) for the normal distribution.

City A: **MENU**; **STAT**; **DIST** (**F5**); **NORM** (**F1**); **Ncd** (**F2**); 29.5 for lower; **EXE**; for upper, any number somewhat greater than the maximum, such as 60 in this case, can be used; **EXE**; 3.54 (the standard deviation calculated previously); **EXE**; 25 (the mean calculated previously); **EXE**; **EXE** or **CALC** (**F1**). The calculated probability is 0.10183. Notice how well this approximates the actual binomial probability calculated in the second method (0.10132).

For Town B, the probability is reported to be 0.25154, which compares favorably to the value calculated from the binomial distribution (0.25173). Once again, the probability is greater for Town B than for City A.

NOTE: Once it has been established that the binomial distribution can be approximated with a normal distribution, some would suggest that the 60% of births be converted to z-scores and the probabilities to the right of those z-scores be compared. This approach does not yield probabilities that are as close to those of the binomial distribution as the process described. For City A, 30 becomes a z-score of approximately 1.41 and the probability of a z-score ≥ 1.41 is 0.079269. Using the z-score for Town B (0.893), the probability is 0.18592. Though these approximations are not as accurate, the conclusion that 60% or more girl births is more likely for Town B than City A is nonetheless unchanged.

