MDM4U: Mathematics of Data Management

Being Discrete

Using the Normal Distribution with Discrete Data

J. Garvin

Probabilities for Continuous Data

Recap

For a normal distribution with $\mu = 28$ and $\sigma = 1.7$, determine $P(x \geq 29.5)$.

$z = \dfrac{29.5 - 28}{1.7} \approx 0.88$.

When $z \approx 0.88$, $P(x \leq 29.5) \approx 0.8106$.

Thus, $P(x \geq 29.5) \approx 1 - 0.8106 \approx 0.1894$.

Probabilities for Discrete Data

While the individual probabilities for continuous data are zero, we can assign non-zero probabilities to discrete data that follow a normal distribution.

For example, assume that a quality-control specialist finds that the number of candies randomly dispensed from a machine is normally distributed.

Since the number of candies dispensed will always be an integer, we can separate the values using the midpoints. Thus, 30 candies might be represented on the normal curve as all values between 29.5 and 30.5.

This process is known as *continuity correction*.

This only applies if the data follow a normal distribution. For other cases, discrete probability distributions (e.g. binomial, hypergeometric) are probably be more appropriate.

Probabilities for Discrete Data

Example

If the mean number of candies dispensed by the machine is 24, with a standard deviation of 3 candies, what is the probability that the machine will dispense at most 28 candies?

Since we want a maximum of 28 candies, we want the area to the left of the curve.

28 candies is represented by the range $27.5 - 28.5$, so use the upper bound, 28.5, to calculate the z-score.

$z = \dfrac{28.5 - 24}{3} = 1.50$.

When $z = 1.50$, $P(x \leq 28.5) \approx 0.9332$.

Therefore, the probability of dispensing at most 28 candies is approximately 0.9332.

Probabilities for Discrete Data

Example

What is the probability of dispensing at least 30 candies?

This time, we want the area to the right of the curve.

30 candies is represented by the range $29.5 - 30.5$, so use the lower bound for the z-score.

$z = \dfrac{29.5 - 24}{3} \approx 1.83$.

When $z \approx 1.83$, $P(x \leq 29.5) \approx 0.9664$.

Therefore, the probability of dispensing at least 30 candies is $1 - 0.9664 \approx 0.0336$.

Probabilities for Discrete Data

Example

What is the probability of dispensing exactly 25 candies?

This time, we want the interval $24.5 - 25.5$.

For $x = 24.5$, $z = \dfrac{24.5 - 24}{3} \approx 0.17$, and for $x = 25.5$, $z = \dfrac{25.5 - 24}{3} = 0.5$.

The corresponding probabilities are 0.5675 and 0.6915.

Therefore, the probability of dispensing exactly 25 candies is $0.6915 - 0.5675 \approx 0.1240$.

Probabilities for Discrete Data

Example

In a sufficient number of tests of the machine, 40% of all tests should result in a number of candies below or equal to some value. How many candies is this?

Perform a reverse-lookup, using a probability of 40%.

The closest probability is 0.4013, corresponding to a z-score of -0.25.

Thus, $-0.25 = \frac{x - 24}{3}$, or $x = -0.25 \times 3 + 24 \approx 23.25$.

This value falls in the interval $22.5 - 23.5$, representing 23 candies. Therefore, 40% of tests should be expected to result in 23 candies or fewer.

Questions?