Probability and Statistics (IT302)

14th September 2020 Monday 09:45 AM - 10:15AM

Class 15

Variability of Continuous Observations about the Mean

If a random variable has a small variance or standard deviation, we would expect most of the values to be grouped around the mean. Therefore, the probability that the random variable assumes a value within a certain interval about the mean is greater than for a similar random variable with a larger standard deviation.

If we think of probability in terms of area, we would expect a continuous distribution with a large value of σ to indicate a greater variability, and therefore we should expect the area to be more spread out, as in Figure 4.2(a). A distribution with a small standard deviation should have most of its area close to μ, as in Figure 4.2(b).

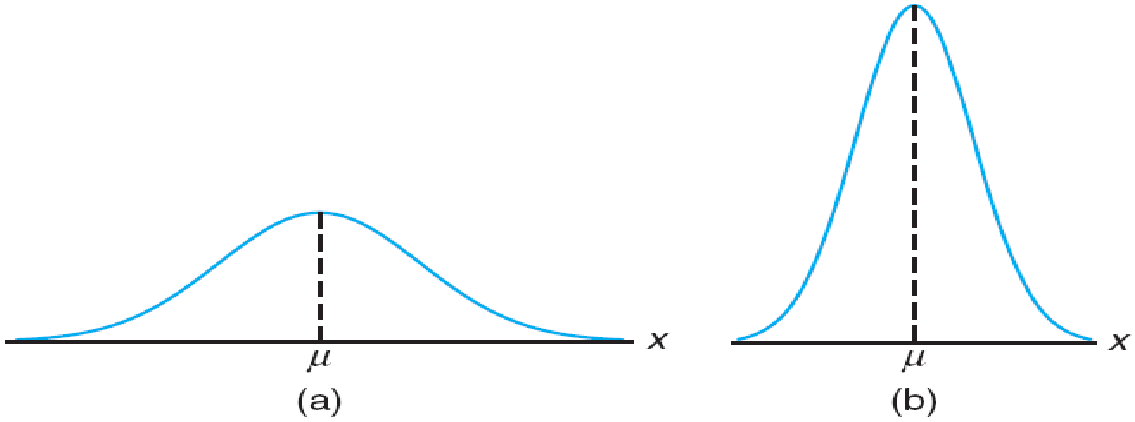

Figure 4.2: Variability of continuous observations about the mean.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Variability of Discrete Observations about the Mean

We can argue the same way for a discrete distribution. The area in the probability histogram in Figure 4.3(b) is spread out much more than that in Figure 4.3(a) indicating a more variable distribution of measurements or outcomes

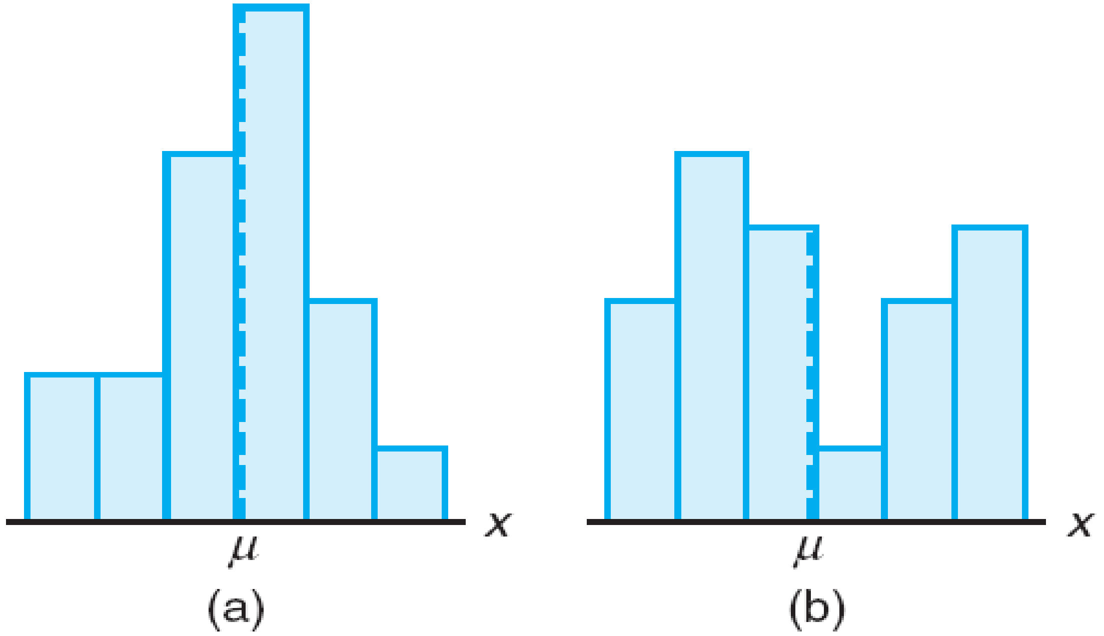

Figure 4.3: Variability of discrete observations about the mean.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Variability of Continuous Observations about the Mean

The Russian mathematician P. L. Chebyshev (1821–1894) discovered that the fraction of the area between any two values symmetric about the mean is related to the standard deviation. Since the area under a probability distribution curve or in a probability histogram adds to 1, the area between any two numbers is the probability of the random variable assuming a value between these numbers.

The following theorem, due to Chebyshev, gives a conservative estimate of the probability that a random variable assumes a value within k standard deviations of its mean for any real number k.

Chebyshev's Theorem

Theorem 4.10: **(Chebyshev's Theorem)** The probability that any random variable X will assume a value within k standard deviations of the mean is at least $1 - 1/k^2$. That is,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

For $k = 2$, the theorem states that the random variable X has a probability of at least $1 - 1/2^2 = 3/4$ of falling within two standard deviations of the mean. That is, three-fourths or more of the observations of any distribution lie in the interval $\mu \pm 2\sigma$. Similarly, the theorem says that at least eight-ninths of the observations of any distribution fall in the interval $\mu \pm 3\sigma$.

Example 4.27

Example 4.27: A random variable X has a mean $\mu = 8$, a variance $\sigma^2 = 9$, and an unknown probability distribution. Find

(a) $P(-4 < X < 20)$,

(b) $P(|X - 8| \geq 6)$.

Solution:

(a) $P(-4 < X < 20) = P[8 - (4)(3) < X < 8 + (4)(3)] \geq \frac{15}{16}$.

(b) $P(|X - 8| \geq 6) = 1 - P(|X - 8| < 6) = 1 - P(-6 < X - 8 < 6)$

$$= 1 - P[8 - (2)(3) < X < 8 + (2)(3)] \leq \frac{1}{4}.$$

Exercise 4.57

Let X be a random variable with the following probability distribution:

x	-3	6	9
$f(x)$	1/6	½	1/3

Find $E(X)$ and $E(X^2)$ and then, using these values, evaluate $E[(2X + 1)^2]$.

Solution:

$$E(X) = (-3)(1/6) + (6)(1/2) + (9)(1/3) = 11/2,$$

$$E(X^2) = (-3)^2 (1/6) + (6)^2(1/2) + (9)^2(1/3) = 93/2. \text{ So,}$$

$$E[(2X + 1)^2] = 4E(X^2) + 4E(X) + 1 = (4)(93/2) + (4)(11/2) + 1 = 209.$$

Exercise 4.65

Let X represent the number that occurs when a red die is tossed and Y the number that occurs when a green die is tossed. Find

(a) $E(X + Y)$; (b) $E(X - Y)$; (c) $E(XY)$.

Solution:

It is easy to see that the expectations of X and Y are both 3.5. So,

(a) $E(X + Y) = E(X) + E(Y) = 3.5 + 3.5 = 7.0.$

(b) $E(X - Y) = E(X) - E(Y) = 0.$

(c) $E(XY) = E(X)E(Y) = (3.5)(3.5) = 12.25.$

Exercise 4.66

Let X represent the number that occurs when a green die is tossed and Y the number that occurs when a red die is tossed. Find the variance of the random variable
(a) $2X - Y$; (b) $X + 3Y - 5$.

Solution:

$$\mu X = \mu Y = 3.5. \ \sigma^2_X = \sigma^2_Y = [(1)^2 + (2)^2 + \cdot\cdot\cdot + (6)^2](1/6) - (3.5)^2 = 35/12 .$$

(a) $\sigma_{2X-Y} = 4\sigma^2_X + \sigma^2 Y = 175/12$;
(b) $\sigma_{X+3Y-5} = \sigma^2_X + 9\sigma^2_Y = 175/6$.

Exercise 4.75

An electrical firm manufactures a 100-watt light bulb, which, according to specifications written on the package, has a mean life of 900 hours with a standard deviation of 50 hours. At most, what percentage of the bulbs fail to last even 700 hours? Assume that the distribution is symmetric about the mean.

Solution:

$\mu = 900$ hours and $\sigma = 50$ hours. Solving $\mu - k\sigma = 700$ we obtain $k = 4$.

So, using Chebyshev's theorem with $P(\mu-4\sigma < X < \mu+4\sigma) \geq 1-1/4^2 = 0.9375$, we obtain $P(700 < X < 1100) \geq 0.9375$. Therefore, $P(X \leq 700) \leq 0.03125$.