EE 464

Spring 2003

Lecture Notes Part 9d

Christopher Wayne Walker, Ph.D.

9.3 Variance

The variance measure gives us an indication of the spread of the data about its mean.

Notation: The variance of X is written $Var(X)$ or σ_X^2 or σ^2.

9.3.1 Discrete Case

Definition: The *variance* of X is given by

$$Var(X) = \sum_i (x_i - \mu)^2 f_X(x_i)$$

where $\mu = E(X)$. Thus, $Var(X) = E\left[(X-\mu)^2\right]$.

9.3.2 Continuous Case

Definition: The *variance* of X is given by

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx$$

where $\mu = E(X)$. Again, $Var(X) = E\left[(X-\mu)^2\right]$.

9.4 Examples and Additional Results

Theorem: Let X be binomially distributed with parameters n, p (write $X \sim B(n,p)$). Then $E(X) = np$.

Proof:

i. Direct proof.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

So,

$$E(X) = \sum_{k=0}^{n} k \frac{n!}{(n-k)!k!} \, p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!} \, p^k (1-p)^{n-k}.$$

Let $s = k - 1$. Then

$$\begin{aligned} E(X) &= \sum_{s=0}^{n-1} \frac{n!}{(n-s-1)!s!} \, p^{s+1}(1-p)^{n-s-1} \\ &= \sum_{s=0}^{n-1} n \binom{n-1}{s} p^{s+1}(1-p)^{n-s-1} \\ &= np \sum_{s=0}^{n-1} \binom{n-1}{s} p^s (1-p)^{n-1-s}. \end{aligned}$$

Recall the binomial theorem:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Replace n by $n-1$, let $a = p,\ b = 1-p$ to get

$$1 = [p + (1-p)]^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

Therefore,

$$\sum_{s=0}^{n-1} \binom{n-1}{s} p^s (1-p)^{n-1-s} = 1$$

and thus $E(X) = np$.

ii. Quick proof.
Think of X as the number of successes in n Bernoulli trials. Let X_i be the number of successes on the ith trial. Then,

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p.$$

Hence,

$$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p.$$

Now,

$$X = X_1 + X_2 + \cdots + X_n.$$

So,

$$\begin{aligned} E(X) = E(X_1 + \cdots + X_n) &= E(X_1) + \cdots + E(X_n) \\ &= p + \cdots + p = np. \end{aligned}$$

Example: Say X has *pdf*

$$f_X(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Let $Y = g(X) = 3X + 1$. Find $E(Y)$.

Previously, (see section 9.1.2) we derived

$$f_Y(y) = \begin{cases} \frac{2}{9}(y-1), & 1 < y < 4, \\ 0, & \text{elsewhere.} \end{cases}$$

So,

$$E(Y) = \int_1^4 y \left[\frac{2}{9}(y-1)\right] dy = 3$$

or (without finding the *pdf* of Y),

$$E(Y) = E[g(X)] = \int_0^1 (3x+1)(2x)dx = 3.$$

Theorem: $Var(X) = E\left(X^2\right) - [E(X)]^2$.

Proof:

$$\begin{aligned} Var(X) &= E\left[(X-\mu)^2\right] = E\left[X^2 - 2X\mu + \mu^2\right] \\ &= E\left(X^2\right) - 2\mu E(X) + \mu^2 \\ &= E\left(X^2\right) - [E(X)]^2 \end{aligned}$$

since $\mu = E(X)$.

Properties of Variance:

i. If c is a constant, $Var(X + c) = Var(X)$.

Proof:

$$\begin{aligned} Var(X+c) &= E\left\{[(X+c) - E(X+c)]^2\right\} \\ &= E\left\{[X + c - E(X) - c]^2\right\} \\ &= E\left\{[X - E(X)]^2\right\} \\ &= Var(X). \end{aligned}$$

ii. If c is a constant, $Var(cX) = c^2 Var(X)$.

Proof: Exercise.

Lemma: If X is discrete and takes values $1, 2, 3, \ldots$, then

$$E(X) = \sum_{n=0}^{\infty} P(X > n).$$

Proof: Let $p_i = P(X = i)$. Then

$$\begin{array}{lcllllllllll}
P(X>0) & = & p_1 & + & p_2 & + & p_3 & + & \cdots & + & p_k & + & & \cdots \\
P(X>1) & = & & & p_2 & + & p_3 & + & \cdots & + & p_k & + & & \cdots \\
P(X>2) & = & & & & & p_3 & + & \cdots & + & p_k & + & & \cdots \\
\vdots & & & & & & & & & & & & & \\
P(X>k) & = & & & & & & & & & & & p_{k+1} & + & \cdots \\
\vdots & & & & & & & & & & & & &
\end{array}$$

By summing along the rows we find the total in the array is

$$\sum_{n=0}^{\infty} P(X > n).$$

By summing along the columns we find the total in the array is

$$p_1 + 2p_2 + 3p_3 + \cdots + kp_k + \cdots.$$

Thus equating these last results we get

$$\sum_{n=0}^{\infty} P(X > n) = p_1 + 2p_2 + 3p_3 + \cdots$$

$$= \sum_{k=1}^{\infty} kp_k = E(X).$$