Class Notes 2-25-2019

- **The Variance of a Random Variable:**
The variance captures the amount of variation of X around the mean μ, presumed to be finite. It is defined as

$$\sigma^2 = \text{var}(X) = E[(X-\mu)^2], \qquad \sigma = \sqrt{\text{var}(X)} = \textit{standard deviation} \text{ of } X$$

The variance is easier to deal with mathematically than the more intuitive variation measure *mean absolute deviation* $E[|X-\mu|]$.
In a way, σ is meant to counter the distortion of the square function used in the variance.

- **Variance Computation and Alternate Forms:**
In the discrete case with pmf $p_X(k)$

$$\begin{aligned}\text{var}(X) &= \sum_k (k-\mu)^2 P(X=k) = \sum_k (k-\mu)^2 p_X(k) = \sum_k k^2 p_X(k) - \sum_k 2k\mu p_X(k) + \sum_k \mu^2 p_X(k) \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 = \text{mean square} - \text{squared mean}\end{aligned}$$

In the continuous case with pdf $f_X(x)$

$$\begin{aligned}\text{var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx = \int_{-\infty}^{\infty} x^2 f_X(x)dx - \int_{-\infty}^{\infty} 2x\mu f_X(x)dx + \int_{-\infty}^{\infty} \mu^2 f_X(x)dx \\ &= E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 = \text{mean square} - \text{squared mean}\end{aligned}$$

The alternate expression $E(X^2) - [E(X)]^2$ for $\text{var}(X)$ holds generally.
$\text{var}(X) = 0 \Leftrightarrow P(X=c) = 1$ for some c, in which case $E(X) = c$.

- **Variance Example:**
Let X take on the 2 values ± 1 with equal probability and Y takes on the 2 values ± 100 with equal probability. We have $E(X) = E(Y) = 0$ but

$$\text{var}(X) = E[X^2] - 0^2 = 1 \quad \text{and} \quad \text{var}(Y) = E[Y^2] - 0^2 = 10{,}000 \quad \text{with } \sigma_X = 1 \text{ and } \sigma_Y = 100$$

- **Variance of a Bernoulli Random Variable:**
Let $X \sim \text{Ber}(p)$ then $\text{var}(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p)$

- **Variance of a Binomial Random Variable:**
Let $X \sim \text{Bin}(n,p)$ and $g(X) = X^2 = X(X-1) + X$. Using $k(k-1)\binom{n}{k} = \binom{n-2}{k-2} n(n-1)$ we get

$$\begin{aligned}E(X^2) &= \sum_{k=0}^{n} [k(k-1)+k] \binom{n}{k} p^k (1-p^{n-k} = \sum_{k=0}^{n} k(k-1) \binom{n}{k} p^k (1-p^{n-k} + np \\ &= p^2 \sum_{k=2}^{n} n(n-1) \binom{n-2}{k-2} p^{k-2} (1-p^{n-2-(k-2)} + np \quad \text{substitute } k-2 = \ell \\ &= p^2 n(n-1) \sum_{\ell=0}^{n-2} \binom{n-2}{\ell} p^{\ell} (1-p^{n-2-\ell} + np = p^2 n(n-1) + np = n^2p^2 - np^2 + np\end{aligned}$$

Thus $E(X^2) - [E(X)]^2 = np(1-p) = \text{var}(X)$.
If we view $X = X_1 + \ldots + X_n$ with $X_i \sim \text{Ber}(p)$ mutually independent we see that

$$\text{var}(X) = np(1-p) = \text{var}(X_1) + \ldots + \text{var}(X_n) = p(1-p) + \ldots + p(1-p)$$

The variance of a sum of independent random variables is the sum of the variances of the individual summands. This will be seen to hold generally later.

- **Variance of $X \sim \text{Unif}[a,b]$ with $a < b$**

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \left. \frac{x^3}{3} \right|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ba + a^2}{3}$$

$$\text{var}(X) = \frac{b^2 + ba + a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{4b^2 + 4ba + 4a^2}{12} - \frac{3a^2 + 6ab + 3b^2}{12} = \frac{(b-a)^2}{12}$$

- **Variance of $X \sim \text{Geo}(p)$:**

$$\begin{aligned} E(X^2) &= E[X(X-1) + X] = E[X(X-1)] + E(X) = \sum_{k=2}^{\infty} k(k-1)pq^{k-1} + \frac{1}{p} \qquad q = 1-p \\ &= pq \sum_{k=2}^{\infty} k(k-1)q^{k-2} + \frac{1}{p} = pq \sum_{k=0}^{\infty} \frac{d^2}{dq^2} q^k + \frac{1}{p} = pq \frac{d^2}{dq^2} \sum_{k=0}^{\infty} q^k + \frac{1}{p} = pq \frac{d^2}{dq^2} \frac{1}{1-q} + \frac{1}{p} \\ &= pq \frac{2}{(1-q)^3} + \frac{1}{p} = \frac{1+q}{p^2} \quad \text{and thus} \quad \text{var}(X) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2} \end{aligned}$$

which has some intuitive appeal for p close to 0 or 1.

- **Expectation and Variance of $aX + b$**
Here a and b are given constant values in $\mathbb{R}$.

$$E(aX + b) = aE(X) + b \qquad \text{and} \qquad \text{var}(aX + b) = a^2 \text{var}(X)$$

- **Binomial Example:**
Let $Z \sim \text{Bin}(10, 1/5)$, find $E(3Z + 2)$ and $\text{var}(3Z + 2)$.

$$E(Z) = 10 \cdot \frac{1}{5} = 2 \quad \text{and} \quad \text{var}(Z) = 10 \cdot \frac{1}{5} \cdot \frac{4}{5} = \frac{8}{5}$$

$$\Rightarrow \quad E(3Z + 2) = 3 \cdot 2 + 2 = 8 \quad \text{and} \quad \text{var}(3Z + 2) = 3^2 \cdot \frac{8}{5} = \frac{72}{5}$$

- **Linear Combinations of Higher Moments:**

$$E\left[\sum_{k=0}^{n} a_k X^k\right] = \sum_{k=0}^{n} a_k E\left[X^k\right] \quad \text{for constants } a_0, a_1, \ldots, a_n \in \mathbb{R}$$

- **Prediction Mean Squared Error:**
  Suppose you will be observing a random variable X and you would want to predict its value by some number a. Further assume that we measure the goodness of that prediction by the mean squared error MSE= $E(X-a)^2$. Which a gives you the smallest MSE?
  Let $\mu = E(X)$ and we have

  $$E(X-a)^2 = E(X-\mu+\mu-a)^2 = E(X-\mu)^2+2(\mu-a)E(X-\mu)+(\mu-a)^2 = E(X-\mu)^2+(\mu-a)^2$$

  which is smallest for $a = \mu$.
  On the other hand $E|X-a|$ is minimized by $a = m =$ the median of X (HW6).