[2.4] Some properties of expected values

Random Variable X. Consider $Y = u(X)$

(A function of a Random Variable is another random variable.

Thm 2.4.1 If X is a Random variable with pdf $f(x)$ and $u(x)$ is a real valued function whose domain includes all the possible values of X, then

$$E[u(x)] = \begin{cases} \int_{-\infty}^{\infty} u(x) f(x)\,dx & \text{, if } X \text{ is continuous} \\ \sum_{\text{all } x} u(x) f(x) & \text{, if } X \text{ is discrete.} \end{cases}$$

Thm 2.4.2 Linearity of the Expectation Operator.

If X is a R.V. with pdf $f(x)$, a, b are constants, and $g(x)$ & $h(x)$ are real-valued functions whose domains include all the possible values of X, then

$$E[a g(x) + b h(x)] = a E g(x) + b E h(x)$$
$$= a E[g(x)] + b E[h(x)]$$

DEF 2.4.1 The variance of a R.V. X is given by

$$\text{var}(X) = E(X-\mu)^2 = E[(X-\mu)^2]$$

Note: σ^2, σ_X^2, and $V(X)$ are all common notations for $\text{var}(X)$. Further $\sigma = \sqrt{\text{var}(X)}$ & σ_X is called the standard deviation.

Corollary

$Var(X) = E(X^2) - \mu^2$, where $\mu = E(X)$.

Proof: $Var(X) = E(X-\mu)^2$ (by def)

$$= E\left[X^2 - 2X\mu + \mu^2\right]$$

$$= E(X^2) - 2\mu \underbrace{E(X)}_{\mu} + \mu^2$$

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

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

$$= E(X^2) - \left[E(X)\right]^2$$

Note: The variance (or std. dev) provides a measure of the "spread."

DEF: The k^{th} moment about the origin of a R.V. X is

$$\mu_k' = E(X^k)$$

and the k^{th} moment about the mean is

$$\mu_k = E\left[(X-\mu)^k\right] = E(X-\mu)^k$$

Note: The mean μ is the first moment about the origin $\mu = \mu_1'$

The variance σ^2 is the second moment about the mean $\sigma^2 = E(X-\mu)^2 = \mu_2$

$\boxed{\text{Thm 7.4.4}}$ If X is a R.V. and a, b are const. then

$$Var(aX+b) = a^2 \, Var(X)$$

Note: the mean absolute deviation is defined as

$$E|X-\mu| \quad \text{(MAD)}$$

Thm 2.4.5 If a dist is symmetric about mean μ, then the third moment about the mean is 0. (e.g. $\mu_3=0$)

Thm. 2.4.6 If X is a R.V. $u(X)$ is a non-neg. real valued function, then for any positive constant $c>0$,

$$P[u(X)\geq c] \leq \frac{E[u(X)]}{c}$$

Markov inequality set $u(x)=|X|^r$, $r>0$

then thm 2.4.6 gives

$$P[|X|\geq c] \leq \frac{E[|X|^r]}{c^r}$$

Thm 2.4.7 Chebychev's Inequality

If X is R.V with mean μ and variance σ^2, then for any $k>0$,

$$P[|X-\mu|\geq k\sigma] \leq \frac{1}{k^2}$$

proof: $u(x)=(X-\mu)^2$

An alternative form is

$$P[|X-\mu|<k\sigma] \geq 1-\frac{1}{k^2}$$

This is the first time you are introduced to the Degenerate Distribution. It concentrates ALL its probability on one value, mu

2.4 Cont.

Thm 2.4.8 Let $\mu = E(X)$ and $\sigma^2 = Var(X)$

If $\sigma^2 = 0$, then $P[X=\mu] = 1$

Proof. If $x \neq \mu$ for some observed value x, then $|x-\mu| \geq \frac{1}{i}$ for some integer $i \geq 1$ and

Thus, $[X \neq \mu] = \bigcup_{i=1}^{\infty} \left[|X-\mu| \geq \frac{1}{i}\right]$

Using Boole's Inequality,

$$P[X \neq \mu] \leq \sum_{i=1}^{\infty} \underbrace{P\left[|X-\mu| \geq \frac{1}{i}\right]}_{\text{Chebychev's Inequality}} \leq \sum_{i=1}^{\infty} i^2 \underbrace{\sigma^2}_{=0} = 0 \Rightarrow P[X \neq \mu] \leq 0$$

$$\Rightarrow P[X \neq \mu] = 0$$

$P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$ $\quad k\sigma = \frac{1}{i}$, $k = \frac{1}{i\sigma} \Rightarrow \frac{1}{k} = \frac{1}{\ldots} = i^2\sigma^2$

So $P[X=\mu] = 1 - P[X \neq \mu]$
$= 1$

Note that this dist is called a degenerate dist (A dist that concentrates all the prob at one pt)

Approximate Mean and Variance

If a function of a Random Var, say $H(X)$ can be expanded in a Taylor series, then an expression for the approx mean and variance can be obtained in terms of the mean & var of X.

Define $\mu = E(x)$.

$$(*)\quad H(x) \overset{\text{approx equal to}}{\doteq} H(\mu) + H'(\mu)(x-\mu) + \frac{1}{2}H''(\mu)(x-\mu)^2$$

$$E[H(x)] \doteq \underbrace{E[H(\mu)]}_{H(\mu)} + H'(\mu)\underbrace{[E(x) - \underset{\mu}{E(\mu)}]}_{\mu-\mu=0} + \frac{1}{2}H''(\mu)\underbrace{E(x-\mu)^2}_{Var(x)=\sigma^2}$$

$$\boxed{E[H(x)] \doteq H(\mu) + \frac{1}{2}H''(\mu)\sigma^2}$$

$$Var(ax+b) = a^2 Var(x)$$

$$\boxed{Var(H(x)) = Var\left(\underbrace{H(\mu)}_{b} + \underset{a}{H'(\mu)}(x-\mu)\right) = [H'(\mu)]^2\sigma^2}$$

The accuracy of these approx depends primarily on the nature of $H(x)$ as well as the amount of variability in X.