Lecture 5

Expectations continued.

Theorem. For $g: \mathbb{R} \to \mathbb{R}$

If X is a discrete r.v. then

$$E[g(X)] = \sum_x g(x) p_X(x)$$

If X is a continuous r.v.

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$$

Proof. Discrete case: $X: x_1, x_2, \ldots, x_n$

$Y = g(X): y_1, y_2, \ldots y_m$, $m \leq n$

show $E[Y] = \sum_x g(x_i) p(x_i)$

$$P(g(X) = y_i) = \underbrace{\sum_{x_j : g(x_j) = y_i} p(x_j)}_{} = p^*(y_i)$$

pmf for Y

$$E[Y] = \sum_{i=1}^{m} y_i p^*(y_i) = \sum_{i=1}^{m} y_i \sum_{x_j : g(x_j) = y_i} p(x_j)$$

$$= \sum_{i=1}^{m} \sum_{x_j : g(x_j) = y_i} y_i p(x_j) = \sum_{j=1}^{m} g(x_j) p(x_j)$$

$\boxed{\Xi}$

Ex. $X: \pm 1, \pm 2, \pm 3$, $Y = X^2$

$Y: 1, 4, 9$ $\quad n = 6, m = 3$

$$E(Y) = E(X^2) = \sum_x x^2 p_X(x) = \big(1 \cdot p(x_1) + 1 \cdot p(x_2)\big) + \big(4p(x_3) + 4p(x_4)\big) + \big(9p(x_5) + 9p(x_6)\big)$$

Ex.

① $X \sim Unif(0,1)$, $Y = X^2$

$$E[Y] = \int_0^1 x^2 \, 1 \, dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

② $X \sim Poisson(\lambda)$, $Y = e^X$

$$E(Y) = \sum_{x=0}^{\infty} e^x \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \underbrace{\sum_{x=0}^{\infty} \frac{(e\lambda)^x}{x!}}$$

$$= e^{-\lambda + \lambda e} = e^{\lambda(e-1)}$$

Properties of Expectation.

Let X, Y be r.v's and $a, b \in \mathbb{R}$

1. $E(a) = a$

$$E(a) = \sum_{p_i} a \, p_i = a \sum p_i = a$$

2. $E(aX + bY) = aE(X) + bE(Y)$

Proof: later (joint dist'n)

3. If X is a non-negative r.v., then $E(X) = 0$ iff $X = 0$ with probability 1.

$\Leftarrow$ $P(X=0) = 1$, $E(X) = 0 \cdot 1 = 0$

4. If X is a non-negative r.v., then $E(X) \geq 0$. $\leftarrow$ def

$\Rightarrow$ $E(X) = 0$. Assume $\exists \, w_0$ s.t. $X(w_0) > 0$ with $p > 0$ $\Rightarrow$ $E(X) = \sum_{w} X(w) P \neq 0$

$\rightarrow \leftarrow$

Moments.

Def. The **k^{th} moment** of a distribution is $E(X^k)$.

Ex. 1. $X \sim Unif(0,1)$

$$E(X^2) = \frac{1}{3}$$

2. $X \sim Geom(p)$ $\qquad E(X) = \frac{1}{p}$

$$E(X^2) = \sum_{x=1}^{\infty} x^2 p q^{x-1} \qquad \frac{2q}{p^2} + \frac{1}{p}$$

$$E[X^2] = E[X(X-1) + X] = E[X(X-1)] + E[X]$$

$$E[X(X-1)] = \sum_{x=1}^{\infty} x(x-1) p q^{x-1} = pq \sum_{x=2}^{\infty} x(x-1) q^{x-2}$$

$$= pq \frac{d^2}{dq^2} \sum_{x=2}^{\infty} q^x = pq \frac{d^2}{dq^2}\left[\frac{1}{1-q} - q - 1\right]$$

$$= pq\left[\frac{1}{(1-q)^2} - 1\right]' = pq \frac{2}{(1-q)^3} = \frac{2q}{p^2}$$

$$\boxed{E(X^2) = \frac{q+1}{p^2}}$$

Variance.

Def. The **variance** of a r.v. X is

$$Var(X) = E[(X - E(X))^2] = E[(X-\mu)^2]$$

Claim: $Var(X) = E(X^2) - E(X)^2 = E(X^2) - \mu^2$

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

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

Def. The standard deviation of a r.v. X is denoted by $\sigma_X = \sqrt{Var(X)}$

Properties of Variance.

Let X, Y be r.v's and $a, b \in \mathbb{R}$.

1. $Var(a) = 0$

   $E[(a-a)^2] = 0$

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

   $E[[aX+b-a\mu-b]^2] = E[a^2(X-\mu)^2] = a^2 E[(X-\mu)^2]$
   $= a^2 Var(X)$

3. $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\, E((X-E(X))(Y-E(Y)))$

   shall prove it later

4. $Var(X) \geq 0$

   $E[\underbrace{(X-\mu)^2}_{\geq 0}] \geq 0$

5. $Var(X) = 0$ iff $X = E(X)$ with probability 1.

   $\Leftarrow$ $P(X=\mu) = 1 \Leftrightarrow P((X-\mu)^2 = 0) = 1$
   $Var(X) = E[(X-\mu)^2] = 0$

   $\Rightarrow$ $Var(X) = 0$. Assume $\exists\, \omega_0$ s.t. $X(\omega_0) \neq \mu$ with $p > 0$
   $E[(X(\omega) - \mu)^2] \neq 0 \longrightarrow \leftarrow$

   ∎

Ex. 1. $X \sim Unif(0,1)$

$$E(X) = \frac{1}{2}$$

$$E(X^2) = \frac{1}{3}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

2. $X \sim Geom(p)$

$$E(X) = \frac{1}{p}$$

$$E(X^2) = \frac{1+q}{p^2}$$

$$Var(X) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

3. $X \sim Bernoulli(p)$

$$E[X] = p$$

$$E(X^2) = 1^2 \cdot p + 0^2 (1-p) = p$$

$$Var(X) = p - p^2 = pq$$

4. $X \sim Gamma(\alpha, \lambda)$

$$Var(X) = \frac{\alpha}{\lambda^2}$$

$X \sim Gamma(\alpha, \beta)$

$$Var(X) = \alpha\beta^2$$

Functions of Random Variables.

Suppose we know the distribution of a r.v. X. We want to find the distribution of $Y = h(X)$.

Discrete case:

$$P_Y(y) = P(Y=y) = P(h(X)=y)$$
$$= P(X = h^{-1}(Y)) = P_X(h^{-1}(y))$$

Ex. $Y = aX + b$

$$P(Y=y) = P(aX+b=y) = P\left(X = \frac{1}{a}(y-b)\right) \quad a \neq 0$$
$$= P_X\left(\frac{1}{a}(y-b)\right)$$

$Y = X^2$

$$P(Y=y) = P(X^2=y) = \begin{cases} P(X=\sqrt{y}) + P(X=-\sqrt{y}), & y>0 \\ P(X=0), & y=0 \\ 0 & y<0 \end{cases}$$

Continuous case:

$X \sim Unif(0,1)$, $Y = X^2$ $\quad F_X(x) = \begin{cases} 0, & x<0 \\ x, & x \in [0,1] \\ 1, & x>1 \end{cases}$

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y})$$
$$= F_X(\sqrt{y}) = \begin{cases} 0, & y<0 \\ \sqrt{y}, & 0 \le y \le 1 \\ 1, & y>1 \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}}, & 0 \le y \le 1 \\ 0, & OW \end{cases}$$

$$E(X^2) = \int_0^1 y \cdot \frac{1}{2\sqrt{y}} \, dy = \frac{1}{3}$$

Q: Can we find a general rule for densities so that we don't have to look at cdf?

$$Y = h(X)$$

$$F_Y(y) = P(h(X) \le y) = P(X \le h^{-1}(y))$$

$$= F_X(h^{-1}(y))$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(h^{-1}(y)) \cdot \frac{d}{dy}[h^{-1}(y)]$$

$$= f_X(h^{-1}(y)) \cdot \frac{d}{dy}[h^{-1}(y)]$$

Theorem. If X is a continuous r.v. with density $f_X(x)$ and h is strictly increasing and differentiable function from $\mathbb{R} \to \mathbb{R}$ then $Y = h(X)$ has density

$$f_Y(y) = f_X(h^{-1}(y)) \frac{d}{dy}[h^{-1}(y)] \quad , y \in \mathbb{R}$$

Proof: see above

$$P(h(X) \le y) = P(h^{-1}(h(X)) \le h^{-1}(y))$$

$$= P(X \le h^{-1}(y))$$

Theorem. If X is a continuous r.v. with density $f_X(x)$ and h is strictly decreasing and differentiable function from $\mathbb{R} \to \mathbb{R}$ then $Y = h(X)$ has density

$$f_Y(y) = -f_X(h^{-1}(y)) \frac{d}{dy}[h^{-1}(y)] \text{ for } y \in \mathbb{R}$$

Proof:

$h^{-1}(y) \searrow$ means if $y_1 < y_2$
$h^{-1}(y_1) > h^{-1}(y_2)$

$$F_Y(y) = P(Y \leq y) = P(h(X) \leq y)$$
$$= P(h^{-1}(h(X)) \geq h^{-1}(y)) = P(X \geq h^{-1}(y))$$
$$= 1 - P(X \leq h^{-1}(y)) = 1 - F_X(h^{-1}(y))$$
$$f_Y(y) = -f_X[h^{-1}(y)] \cdot \frac{d}{dy} h^{-1}(y)$$

In general, $f_Y(y) = f_X(h^{-1}(y)) \left| \frac{d}{dy}(h^{-1}(y)) \right|$

Ex. X has a density $f_X(x) = \begin{cases} \frac{x^3}{4}, & 0 < x < 2 \\ 0, & \text{ow} \end{cases}$

$Y = X^6$, find $f_Y(y)$.

$h^{-1}(y) = \sqrt[6]{y}$

$$f_Y(y) = \frac{(\sqrt[6]{y})^3}{4} \cdot \frac{d}{dy} \sqrt[6]{y} = \frac{\sqrt{y}}{4} \frac{1}{6} y^{-5/6} =$$
$$= \frac{1}{24} y^{-1/3}, \quad 0 < y < 64$$

Indicator Functions and Random Variables.

<u>Def</u>. Let A be a set of real numbers. The indicator function for A is defined by

$$I_A(x) = I\{x \in A\} = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Some properties:

① $I_A(x) I_B(x) = I_{A \cap B}(x)$

② $g(x) I_A(x) = \begin{cases} g(x), & x \in A \\ 0, & x \notin A \end{cases}$

③ $[I_A(x)]^n = I_A(x)$ for any $n > 0$

④ If A is event, $I_A(x)$ – random variable

$E(I_A)$

$Var(I_A)$

<u>Ex</u>. (1) $X \sim p_X(x) = \begin{cases} \frac{x}{6}, & x = 1,2,3 \\ 0, & ow \end{cases}$

$$p_X(x) = \frac{x}{6} I\{x = 1,2,3\} = \frac{x}{6} I_{\{1,2,3\}}(x)$$

(2) $X \sim f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & ow \end{cases}$

$$f_X(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$$

If A is an event, then I_A is a r.v. which is 0 if A does not occur, and 1, if A occurs.

I_A is a Bernoulli r.v.

1101100010111101

$E(I_A)$ = proportion with which A occurs

$$= 1 \cdot P(A) + 0 \cdot P(\bar{A}) = P(A)$$

Ex $X \sim \text{Bernoulli}(p)$

$$E(X) = E(I_A) = P(A) = p$$

$$\text{Var}(I_A) = E(I_A^2) - E(I_A)^2$$

$$= E(I_A) - E(I_A)^2$$

$$= P(A) - P(A)^2$$

$$= P(A)\,P(\bar{A})$$

$$X \sim \text{Bernoulli}(p) \Rightarrow \text{Var}(X) = pq$$