Recap: Function of a random variable

- If X is a random variable and $g : \Re \rightarrow \Re$ is a function, then $Y = g(X)$ is a random variable.
- More formally, Y is a random variable if g is a Borel measurable function.
- We can determine distribution of Y given the function g and the distribution of X

Recap

- Let X be a rv and let $Y = g(X)$.
- The distribution function of Y is given by

$$\begin{aligned} F_Y(y) &= P[g(X) \leq y] \\ &= P[X \in \{z \ : \ g(z) \leq y\}] \end{aligned}$$

- This probability can be obtained from distribution of X.
- We have seen many specific examples of this.

Recap

- Suppose X is a discrete rv with $X \in \{x_1, x_2, \cdots\}$.
- Suppose $Y = g(X)$.
- Then Y is also discrete and $Y \in \{g(x_1), g(x_2), \cdots\}$.
- We can find the pmf of Y as

$$
\begin{aligned}
f_Y(y) &= p[Y = y] = P[g(X) = y] \\
&= P[X \in \{x_i \ : \ g(x_i) = y\}] \\
&= \sum_{\substack{i: \\ g(x_i)=y}} f_X(x_i)
\end{aligned}
$$

Recap

- Let $g : \Re \rightarrow \Re$ be differentiable with $g'(x) > 0, \forall x$ or $g'(x) < 0, \forall x$.
- Let X be a continuous rv and let $Y = g(X)$.
- Then Y is a continuous rv with pdf

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, \ a \leq y \leq b$$

  where $a = \min(g(\infty),\ g(-\infty))$ and $b = \max(g(\infty),\ g(-\infty))$
- This theorem is useful in some cases to find the densities of functions of continuous random variables

Expectation and Moments of a random variable

- We next consider the important notion of expectation of a random variable

Expectation of a discrete rv

- Let X be a discrete rv with $X \in \{x_1, x_2, \cdots\}$
- We define its expectation by

$$E[X] = \sum_i x_i \; f_X(x_i)$$

- Expectation is essentially a weighted average.
- To make the above finite and well defined, we can stipulate the following as condition for existence of expectation

$$\sum_i |x_i| \; f_X(x_i) < \infty$$

Expectation of a Continuous rv

- If X is a continuous random variable with pdf, f_X, we define its expectation as
$$E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$
- Once again we can use the following as condition for existence of expectation
$$\int_{-\infty}^{\infty} |x| \, f_X(x) \, dx < \infty$$
- Sometimes we use the following notation to denote expectation of both kinds of rv
$$E[X] = \int_{-\infty}^{\infty} x \, dF_X(x)$$
- Though we consider only discrete or continuous rv's, expectation is defined for all random variables.

- Let us look at a couple of simple examples.
- Let $X \in \{1, 2, 3, 4, 5, 6\}$ and $f_X(k) = \frac{1}{6},\ 1 \leq k \leq 6$.

$$EX = \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$$

- Let $X \sim U[0,\ 1]$

$$EX = \int_{-\infty}^{\infty} x\ f_X(x)\ dx = \int_0^1 x\ dx = 0.5$$

- When an rv takes only finitely many values or when the pdf is non-zero only on a bounded set, the expectation is always finite.

- The way we have defined existence of expectation, implies that expectation is always finite (when it exists).
- This may be needlessly restrictive in some situations. We redefine it as follows.
- Let X be a non-negative (discrete or continuous) random variable.
- We define its expectation by

$$EX = \sum_i x_i \, f_X(x_i) \quad \text{or} \quad EX = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$

  depending on whether it is discrete or continuous
  (In this course we will consider only discrete or continuous rv's)
- Note that the expectation may be infinite.
- But it always exists for non-negative random variables.

- Now let X be a rv that may not be non-negative.
- We define positive and negative parts of X by

$$X^+ = \begin{cases} X & \text{if } \; X > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$X^- = \begin{cases} -X & \text{if } \; X < 0 \\ 0 & \text{otherwise} \end{cases}$$

- Note that both X^+ and X^- are non-negative. Hence their expectations exist. (Also, $X(\omega) = X^+(\omega) - X^-(\omega), \; \forall \omega$).
- Now we define expectation of X by

$$EX = EX^+ - EX^-, \quad \text{if at least one of them is finite}$$

  Otherwise EX does not exist.
- Now, expectation does not exist only when $EX^+ = EX^- = \infty$

- This is the formal way of defining expectation of a random variable.
- We first note that if $\sum_i |x_i| \ f_X(x_i) < \infty$ then both EX^+ and EX^- would be finite and we can simply take the expectation as $EX = \sum_i x_i \ f_X(x_i)$.
- Also note that if X takes only finitely many values, the above always holds.
- Similar comments apply for a continuous random variable.
- This is what we do in this course because we deal with only discrete and continuous rv's.
- But to get a feel for the more formal definition, we look at a couple of examples.

- Let $X \in \{1, 2, \cdots\}$.
- Suppose $f_X(k) = \frac{C}{k^2}$.
- Since $\sum_k \frac{1}{k^2} < \infty$, we can find C so that $\sum_k f_X(k) = 1$. ($\sum_k \frac{1}{k^2} = \frac{\pi^2}{6}$ and hence $C = \frac{6}{\pi^2}$).
- Hence we get

$$\sum_k |x_k| \; f_X(x_k) = \sum_k x_k \; f_X(x_k) = \sum_k k \; \frac{C}{k^2} = \sum_k \frac{C}{k} = \infty$$

- Here the expectation is infinity.
- But by the formal definition it exists. (Note that here $X^+ = X$ and $X^- = 0$).

- Now suppose X takes values $1, -2, 3, -4, \cdots$ with probabilities $\frac{C}{1^2}$, $\frac{C}{2^2}$, $\frac{C}{3^2}$ and so on.
- Once again $\sum_k |x_k| f_X(x_k) = \infty$.
- But $\sum_k x_k f_X(x_k)$ is an alternating series.
- Here X^+ would take values $2k-1$ with probability $\frac{C}{(2k-1)^2}$, $k = 1, 2, \cdots$
  (and the value 0 with remaining probability).
- Similarly, X^- would take values $2k$ with probability $\frac{C}{(2k)^2}$, $k = 1, 2, \cdots$ (and the value 0 with remaining probability).

$$EX^+ = \sum_k \frac{C}{2k-1} = \infty, \quad \text{and} \quad EX^- = \sum_k \frac{C}{2k} = \infty$$

- Hence EX does not exist.

- Consider a continuous random variable X with pdf

$$f_X(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}, \ -\infty < x < \infty$$

- This is called (standard) Cauchy density. We can verify it integrates to 1

$$\int_{-\infty}^{\infty} \frac{1}{\pi}\,\frac{1}{1+x^2}\,dx = \frac{1}{\pi}\ \tan^{-1}(x)\big|_{-\infty}^{\infty} = \frac{1}{\pi}\left(\frac{\pi}{2} - \frac{-\pi}{2}\right) = 1$$

- What would be EX?

$$EX = \int_{-\infty}^{\infty} x\,\frac{1}{\pi}\,\frac{1}{1+x^2}\,dx \ \stackrel{?}{=} 0 \text{ because } \int_{-a}^{a} \frac{x}{1+x^2} = 0?$$

- The question was

$$EX = \int_{-\infty}^{\infty} x \, \frac{1}{\pi} \, \frac{1}{1+x^2} \, dx \stackrel{?}{=} 0$$

- This depends on the definition of infinite integrals

$$\begin{aligned} \int_{-\infty}^{\infty} g(x) \, dx \quad &\triangleq \quad \lim_{c \to \infty, d \to \infty} \int_{-c}^{d} g(x) \, dx \\ &= \quad \lim_{c \to \infty} \int_{-c}^{0} g(x) \, dx + \lim_{d \to \infty} \int_{0}^{d} g(x) \, dx \end{aligned}$$

This is not same as $\displaystyle \lim_{a \to \infty} \int_{-a}^{a} g(x) \, dx,$

which is known as Cauchy principal value

- Here we have

$$\lim_{c \to \infty} \int_{-c}^{0} \frac{x}{1+x^2} \, dx = -\infty; \quad \lim_{d \to \infty} \int_{0}^{d} \frac{x}{1+x^2} \, dx = \infty$$

- Hence $EX = \int_{-\infty}^{\infty} x \, \frac{1}{\pi} \, \frac{1}{1+x^2} \, dx$ does not exist.
- Essentially, both halves of the integral are infinite and hence we get $\infty - \infty$ type expression which is undefined.
- However, $\lim_{a \to \infty} \int_{-a}^{a} x \, \frac{1}{\pi} \, \frac{1}{1+x^2} \, dx = 0$.

Expectation of a random variable

- Let X be a discrete rv with $X \in \{x_1, x_2, \cdots\}$. Then
$$E[X] = \sum_i x_i \, f_X(x_i)$$
- If X is a continuous random variable with pdf, f_X,
$$E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$
- Sometimes we use the following notation to denote expectation of both kinds of rv
$$E[X] = \int_{-\infty}^{\infty} x \, dF_X(x)$$
- We take the expectation to exist when the sum or integral above is absolutely convergent
- Note that expectation is defined for all random variables
- Let us calculate expectations of some of the standard distributions.

Binary random variable

- Expectation of a binary rv (e.g., Bernoulli):

$$EX = 0 \times f_X(0) + 1 \times f_X(1) = P[X = 1]$$

- Expectation of a binary random variable is same as the probability of the rv taking value 1.
- Thus, for example, $EI_A = P(A)$.

Expectation of Binomial rv

- Let $f_X(k) = {}^nC_k p^k (1-p)^{n-k},\ k = 0, 1, \cdots, n$.

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

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

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

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

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

Expectation of Poisson rv

- $f_X(k) = \frac{\lambda^k}{k!}\, e^{-\lambda},\ k = 0, 1, \cdots$

$$\begin{aligned} EX &= \sum_{k=0}^{\infty} k\, \frac{\lambda^k}{k!}\, e^{-\lambda} \\ &= \lambda \end{aligned}$$

(Left as an exercise for you!)

Expectation of Geometric rv

- $f_X(k) = (1-p)^{k-1}\, p, \quad k = 1, 2, \cdots$

$$EX = \sum_{k=1}^{\infty} k\, (1-p)^{k-1}\, p$$

- We have

$$\sum_{k=1}^{\infty} (1-p)^k = \frac{1-p}{p} = \frac{1}{p} - 1$$

- Term-wise differentiation of the above gives

$$\sum_{k=1}^{\infty} k\, (1-p)^{k-1} = \frac{1}{p^2}$$

- This gives us $EX = \frac{1}{p}$

Expectation of uniform density

- Let $X \sim U[a, b]$. $f_X(x) = \frac{1}{b-a},\ a \leq x \leq b$

$$
\begin{aligned}
EX &= \int_{-\infty}^{\infty} x\ f_X(x)\ dx \\
&= \int_a^b x\ \frac{1}{b-a}\ dx \\
&= \frac{1}{b-a}\ \left.\frac{x^2}{2}\right|_a^b \\
&= \frac{1}{b-a}\ \frac{b^2 - a^2}{2} \\
&= \frac{b+a}{2}
\end{aligned}
$$

Expectation of exponential density

- $f_X(x) = \lambda\, e^{-\lambda x}, \;\; x > 0.$

$$
\begin{aligned}
EX &= \int_0^\infty x\, \lambda\, e^{-\lambda x}\, dx \\
&= x\, \lambda\, \frac{e^{-\lambda x}}{-\lambda}\Big|_0^\infty - \int_0^\infty \lambda\, \frac{e^{-\lambda x}}{-\lambda}\, dx \\
&= \int_0^\infty e^{-\lambda x}\, dx \\
&= \frac{e^{-\lambda x}}{-\lambda}\Big|_0^\infty \\
&= \frac{1}{\lambda}
\end{aligned}
$$

Expectation of Gaussian density

- $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; -\infty < x < \infty$

$$
\begin{aligned}
EX &= \int_{-\infty}^{\infty} x \, \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \\
&\quad \text{make a change of variable } \; y = \frac{x-\mu}{\sigma} \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \, (\sigma y + \mu) e^{-\frac{y^2}{2}} \, dy \\
&= \sigma \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \, e^{-\frac{y^2}{2}} \, dy + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, dy \\
&= \mu
\end{aligned}
$$

Expectation of a function of a random variable

- Let X be a rv and let $Y = g(X)$.
- **Theorem**: $EY = \int y \, dF_Y(y) = \int g(x) \, dF_X(x)$
- That is, if X is discrete, then

$$EY = \sum_j y_j \, f_Y(y_j) = \sum_i g(x_i) f_X(x_i)$$

- If X and Y are continuous

$$EY = \int y \, f_Y(y) \, dy = \int g(x) \, f_X(x) \, dx$$

- This theorem is true for all rv's. But we will prove it in only some special cases.

- **Theorem**: Let $X \in \{x_1, x_2, \cdots x_n\}$ and let $Y = g(X)$. Then

$$EY = \sum_i g(x_i)\ f_X(x_i)$$

- **Proof**: Let $Y \in \{y_1, y_2, \cdots, y_m\}$.
  Each y_j would be equal to $g(x_i)$ for one or more i.
- Let $B_j = \{x_i\ :\ g(x_i) = y_j\}$. Thus,

$$f_Y(y_j) = P[Y = y_j] = P[X \in B_j] = \sum_{\substack{i: \\ x_i \in B_j}} f_X(x_i)$$

- Note that
  - B_j are disjoint
  - each x_i would be in one (and only one) of the B_j

- Now we have

$$
\begin{aligned}
EY &= \sum_{j=1}^{m} y_j f_Y(y_j) \\
&= \sum_{j=1}^{m} y_j \sum_{\substack{i: \\ x_i \in B_j}} f_X(x_i) \\
&= \sum_{j=1}^{m} \sum_{\substack{i: \\ x_i \in B_j}} g(x_i) f_X(x_i) \\
&= \sum_{i=1}^{n} g(x_i) f_X(x_i)
\end{aligned}
$$

  That completes the proof.
- The proof goes through even when X (and Y) take countably infinitely many values (because we assume the expectation sum is absolutely convergent).

- Suppose X is a continuous rv and suppose g is a differentiable function with $g'(x) > 0, \ \forall x$. Let $Y = g(X)$
- Once again we can show $EY = \int g(x) \ f_X(x) \ dx$

$$
\begin{aligned}
EY &= \int_{-\infty}^{\infty} y \ f_Y(y) \ dy \\
&= \int_{g(-\infty)}^{g(\infty)} y \ f_X(g^{-1}(y)) \ \frac{d}{dy} g^{-1}(y) \ dy,
\end{aligned}
$$

change the variable to $x = g^{-1}(y) \ \Rightarrow \ dx = \frac{d}{dy} g^{-1}(y) \ dy$

$$
= \int_{-\infty}^{\infty} g(x) \ f_X(x) \ dx
$$

- We can similarly show this for the case where $g'(x) < 0, \ \forall x$

- We proved the theorem only for discrete rv's and for some restricted case of continuous rv's.
- However, this theorem is true for all random variables.
- Now, for any function, g, we can write

$$E[g(X)] = \sum_i g(x_i) f_X(x_i) \quad \text{or} \quad E[g(X)] = \int_{-\infty}^{\infty} g(x)\ f_X(x)\ dx$$

Some Properties of Expectation

$$E[g(X)] = \sum_i g(x_i) f_X(x_i) \quad \text{or} \quad E[g(X)] = \int_{-\infty}^{\infty} g(x)\ f_X(x)\ dx$$

- If $X \geq 0$ then $EX \geq 0$
- $E[b] = b$ where b is a constant
- $E[ag(X)] = aE[g(X)]$ where a is a constant
- $E[aX + b] = aE[X] + b$ where a, b are constants.
- $E[ag_1(X) + bg_2(X)] = aE[g_1(X)] + bE[g_2(X)]$

- Consider the problem: $\min_c \; E[(X-c)^2]$
- We are asking what is the best constant to approximate a rv with
- We are trying to minimize (weighted) average, over all values X can take, of the square of the error
- We are interested in the best mean-square approximation of X by a constant.

$$E[(X-c)^2] = E[X^2 + c^2 - 2cX] = E[X^2] + c^2 - 2cE[X]$$

- We differentiate this and equate to zero to get the best c

$$2c^* = 2E[X] \;\Rightarrow\; c^* = E[X]$$

- We can derive this in an alternate manner too

$$
\begin{aligned}
E[(X-c)^2] &= E[(X-EX+EX-c)^2] \\
&= E[(X-EX)^2+(EX-c)^2+2(EX-c)(X-EX)] \\
&= E[(X-EX)^2]+(EX-c)^2+2(EX-c)E[(X-EX)] \\
&= E[(X-EX)^2]+(EX-c)^2+2(EX-c)(EX-EX) \\
&= E[(X-EX)^2]+(EX-c)^2 \\
&\geq E[(X-EX)^2]
\end{aligned}
$$

- Thus $E[(X-c)^2] \geq E[(X-EX)^2], \ \forall c$
- So, $E[(X-c)^2]$ is minimized when $c = EX$ and the minimum value is $E[(X-EX)^2]$

Variance of a Random variable

- We define variance of X as $E[(X - EX)^2]$ and denote it as Var(X).
- By definition, Var$(X) \geq 0$.

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

- This also implies: $E[X^2] \geq (EX)^2$

Some properties of variance

- $\text{Var}(X + c) = \text{Var}(X)$ where c is a constant

$$\text{Var}(X+c) = E\left[\{(X+c) - E[X+c]\}^2\right] = E\left[(X - EX)^2\right] = \text{Var}(X)$$

- $\text{Var}(cX) = c^2\text{Var}(X)$ where c is a constant

$$\text{Var}(cX) = E\left[(cX - E[cX])^2\right] = E\left[(cX - cE[X])^2\right] = c^2\text{Var}(X)$$

Variance of uniform rv

- $f_X(x) = \frac{1}{b-a}, \; a \leq x \leq b$

$$
\begin{aligned}
E[X^2] &= \int_a^b x^2 \, \frac{1}{b-a} \, dx \\
&= \frac{1}{b-a} \; \left. \frac{x^3}{3} \right|_a^b \\
&= \frac{1}{b-a} \; \frac{b^3 - a^3}{3} \\
&= \frac{b^2 + ab + a^2}{3}
\end{aligned}
$$

Variance of uniform rv

- We got $E[X^2] = \frac{b^2+ab+a^2}{3}$. Earlier we showed $EX = \frac{b+a}{2}$
- Now we can calculate Var(X) as

$$\begin{aligned}
\text{Var}(X) &= EX^2 - (EX)^2 \\
&= \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} \\
&= \frac{4(b^2+ab+a^2) - 3(b^2+2ab+a^2)}{12} \\
&= \frac{(b^2-2ab+a^2)}{12} \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

Variance of exponential rv

- $f_X(x) = \lambda\, e^{-\lambda x},\ x > 0$

$$\begin{aligned} E[X^2] &= \int_0^\infty x^2\, \lambda\, e^{-\lambda x}\, dx \\ &= x^2\, \lambda\, \frac{e^{-\lambda x}}{-\lambda}\Big|_0^\infty - \int_0^\infty \lambda\, \frac{e^{-\lambda x}}{-\lambda}\, 2x\, dx \\ &= \frac{2}{\lambda} \int_0^\infty x\, \lambda\, e^{-\lambda x}\, dx \\ &= \frac{2}{\lambda^2} \end{aligned}$$

- Hence the variance is now given by

$$\mathsf{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Variance of Gaussian rv

- Let $X \sim \mathcal{N}(0,1)$. That is,
  $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \ -\infty < x < \infty$.
- We know $EX = 0$. Hence $\mathrm{Var}(X) = EX^2$.

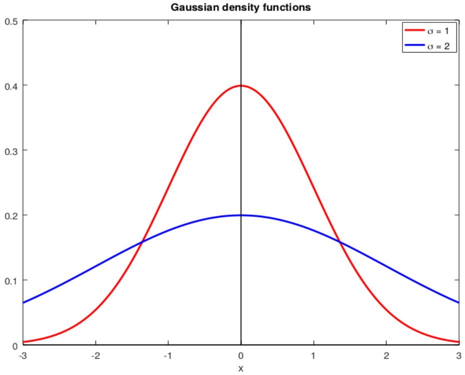
$$\begin{aligned}
\mathrm{Var}(X) &= EX^2 = \int_{-\infty}^{\infty} x^2 \, \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \\
&= \int_{-\infty}^{\infty} x \, \left(x \, \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \, dx \\
&= x \, \left. \frac{-1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \\
&= 1
\end{aligned}$$

- Let $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \; -\infty < x < \infty$.
- Let $g(x) = \sigma x + \mu$ and hence $g^{-1}(y) = \frac{y-\mu}{\sigma}$.
- Take $\sigma > 0$ and $Y = g(X)$. By the theorem,

$$f_Y(y) = \left(\frac{d}{dy} g^{-1}(y)\right) \; f_X(g^{-1}(y)) = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

- Since $Y = \sigma X + \mu$, we get
  - $EY = \sigma EX + \mu = \mu$
  - $\text{Var}(Y) = \sigma^2 \text{Var}(X) = \sigma^2$
- When $Y \sim \mathcal{N}(\mu, \sigma^2)$, $EY = \mu$ and $\text{Var}(Y) = \sigma^2$.

- Here is a plot of Gaussian densities with different variances



Variance of Binomial rv

- $f_X(k) = \frac{n!}{k!(n-k)!}\, p^k\, (1-p)^{n-k},\ k = 0, 1, \cdots, n$
- Here we use the identity, $EX^2 = E[X(X-1)] + EX$

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^{n} k(k-1)\, \frac{n!}{k!(n-k)!}\, p^k\, (1-p)^{n-k} \\ &= \sum_{k=2}^{n} k(k-1)\, \frac{n!}{k!(n-k)!}\, p^k\, (1-p)^{n-k} \end{aligned}$$

$$= \sum_{k=2}^{n} \frac{n(n-1)(n-2)!}{(k-2)!((n-2)-(k-2))!}\, p^2\, p^{k-2}\, (1-p)^{(n-2)-(k-2)}$$

$$= n(n-1)p^2 \sum_{k'=0}^{n-2} \frac{(n-2)!}{k'!((n-2)-k')!}\, p^{k'}\, (1-p)^{(n-2)-k'}$$

$$= n(n-1)p^2$$

- When X is binomial rv, we showed, $E[X(X-1)] = n(n-1)p^2$
- Hence,

$$EX^2 = E[X(X-1)] + EX = n(n-1)p^2 + np = n^2p^2 + np(1-p)$$

- Now we can calculate the variance

$$\text{Var(X)} = EX^2 - (EX)^2 = n^2p^2 + np(1-p) - (np)^2 = np(1-p)$$

Variance of a geometric random variable

- $X \in \{1, 2, \cdots\}$ and $f_X(k) = (1-p)^{k-1}p, \; k = 1, 2, \cdots$
- Here also, it is easier to calculate $E[X(X-1)]$

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

- We know

$$\sum_{k=1}^{\infty}(1-p)^k = \frac{1-p}{p} \;\Rightarrow\; \sum_{k=1}^{\infty} k(k-1)(1-p)^{k-2} = \frac{d^2}{dp^2}\left(\frac{1-p}{p}\right)$$

Now you can compute $E[X(X-1)]$ and hence $E[X^2]$ and hence Var(X) and show it to be equal to $\frac{1-p}{p^2}$.
(Left as an exercise)

moments of a random variable

- We define the k^{th} order moment of a rv, X, by

$$m_k = E[X^k] = \int x^k \, dF_X(x)$$

- $m_1 = EX$ and $m_2 = EX^2$ and so on
- We define the k^{th} central moment of X by

$$s_k = E[(X - EX)^k] = \int (x - EX)^k \, dF_X(x)$$

- $s_1 = 0$ and $s_2 = \text{Var}(X)$.
- Not all moments may exist for a given random variable. (For example, m_1 does not exist for Cauchy rv)

- **Theorem**: If $E\left[|X|^k\right] < \infty$ then $E\left[|X|^s\right] < \infty$ for $0 < s < k$.
- For example, if third order moment exists then so do first and second order moments
- **Proof**: We prove it when X is continuous rv. Proof for discrete case is similar.

$$
\begin{aligned}
E\left[|X|^s\right] &= \int_{-\infty}^{\infty} |x|^s\, f_X(x)\, dx \\
&= \int_{|x|<1} |x|^s\, f_X(x)\, dx + \int_{|x|\geq 1} |x|^s\, f_X(x)\, dx \\
&\leq \int_{|x|<1} f_X(x)\, dx + \int_{|x|\geq 1} |x|^s\, f_X(x) \\
&\leq P[|X|^s < 1] + \int_{|x|\geq 1} |x|^k\, f_X(x) \\
&\qquad \text{since for } |x| \geq 1,\ |x|^s < |x|^k \text{ when } s < k \\
&< \infty \ \ \text{because } E\left[|X|^k\right] = \int_{\infty}^{\infty} |x|^k f_X(x)\, dx < \infty
\end{aligned}
$$