2/21/18

Stat 134 lec 15

Sec 3.4 Discrete Distributions, — allow RV w/ countably many outcomes.

Countable Additivity Axiom

let $A_1, A_2, \dots$ mutually exclusive

then $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Expectation of a discrete RV

$$E(X) = \sum_{\text{all } x} x P(X=x) \quad \text{provided } E(|X|) \text{ is finite } (<\infty)$$

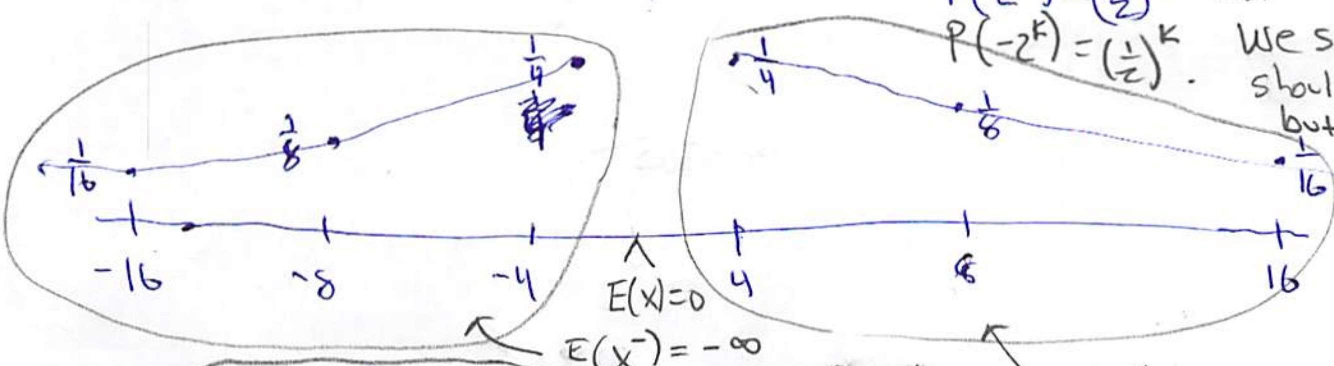
Similarly for $E(g(X))$.

$X^+ = 2^k \; k \ge 2$, $X^- = -2^k \; k \ge 2$, $X = X^- + X^+$

Example where $E(X)$ not defined: let $X = \pm 2^k$, $k \ge 2$

$P(2^k) = \left(\frac{1}{2}\right)^k$ and

$P(-2^k) = \left(\frac{1}{2}\right)^k$. We see $E(X)$ should be zero but it is $-\infty + \infty$ which is not well defined,



$X = X^- + X^+$

$E(X) = E(X^-) + E(X^+)$

$= -\infty + \infty$ ← not even defined.

Geometric distribution

iid Bernoulli (p) trials

X = # trials until the first success.

$X \sim \text{Geom}(p)$

$P(X > k) = q^k$

Use tail sums formula to find the Expectation

$$E(X) = \sum_{k=1}^{\infty} P(X \ge k) = \sum_{k=0}^{\infty} P(X > k) = \sum_{k=0}^{\infty} q^k = \frac{1}{1-q} = \frac{1}{p}$$

Stat 134

Wednesday February 21 2018

1. You have a coin that has probabililty p of landing heads. Let X be the number of coin tosses until you get heads.
   What is $P(X > k)$?

   a q^k where $q = 1 - p$.

   b $q^{k-1}p$

   c $q^{k-1}p(1 + q + q^2 + \dots)$

   d none of the above

Always start w/ poss. values of X? 1, 2, 3, ...

$$P(X>k) = \underset{\substack{\| \\ q^k p}}{P(X=k+1)} + \underset{\substack{\| \\ q^{k+1} p}}{P(X=k+2)} + \cdots$$

$$= q^k p \underbrace{(1 + q + q^2 + \cdots)}_{\substack{\| \\ \frac{1}{1-q} = \frac{1}{p}}} = \boxed{q^k}$$

chance of getting k failures in a row.

Trick to find var(x):

Find $E(X(X-1))$ —— $Var(X) = E(X^2) - E(X)^2$

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

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

$= E(X(X-1)) + \underbrace{E(X)}_{=\frac{1}{p}} - \underbrace{E(X)^2}_{=\frac{1}{p^2}}$

to find $E(X(X-1))$:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

$\frac{d}{dq}$

$$\sum_{k=0}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

$\frac{d}{dq}$

$$\boxed{\sum_{k=0}^{\infty} k(k-1) q^{k-2} = \frac{2}{(1-q)^3} = \frac{2}{p^3}}$$

$$E(X(X-1)) = \sum_{k=0}^{\infty} k(k-1)\underbrace{P(X=k)}_{=q^{k-1}p} = qp\underbrace{\sum_{k=0}^{\infty} k(k-1) q^{k-2}}_{=\frac{2}{p^3}} = \boxed{\frac{2q}{p^2}}$$

$$E(g(X)) = \sum_{\text{all } x} g(x) \cdot P(X=x)$$

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

$$\boxed{SD(X) = \frac{\sqrt{q}}{p}}$$

Note

Confusing! Some books define geometric (p)
as Y = # failures until the first success

0, 1, 2, 3

$Y = X - 1$ ← # trials until the first success.

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

$SD(Y) = SD(X) = \frac{\sqrt{q}}{p}$.

Negative Binomial Dist with param r and p

Generalization of Geom ($r=1$) — # trials until the r^{th} success.

Ex $r=3$

$\underbrace{qqqp}_{W_1}\underbrace{qqqp}_{W_2}\underbrace{p}_{W_3}$

$T_r = W_1 + W_2 + \cdots + W_r$ where W_k = # tosses after the $(k-1)^{st}$ success including the k^{th} success.

$W_1, W_2, \ldots$ are iid Geom (p)

$E(T_r) = \boxed{r \cdot \frac{1}{p}}$

$Var(T_r) = r \cdot \frac{q}{p^2} \Rightarrow \boxed{SD(T_r) = \frac{\sqrt{rq}}{p}}$

Dist of T_r :

poss values $r, r+1, \ldots$

$P(T_r = k)$ = Chance it takes me k trials to get the r^{th} success.

$r-1$ successes in $k-1$ trials

$\overbrace{\qquad\qquad\qquad\qquad}\ p$

$\binom{k-1}{r-1} p^{r-1} q^{k-r}$

$$\boxed{P(T_r = k) = \binom{k-1}{r-1} p^{r-1} q^{k-r} \cdot p}$$

ex Coupon collector's Problem

Have a collection of boxes each containing a coupon, n diff coupons. Each box is equally likely to contain any coupon indep of the other boxes.

Find $E(X)$ where X = # boxes needed to get all n diff coupons.

X_1 = # boxes to 1^{st} coupon

X_2 = # " " 2^{nd} diff coupon after 1^{st} coupon.

X_3 = # " " 3^{rd} different coupon after 2^{nd} coupon.

$\vdots$

X_n

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

$X_1 \sim Geom(1)$

$X_2 \sim Geom\left(\frac{n-1}{n}\right)$

$\vdots$

$X_n \sim Geom\left(\frac{1}{n}\right)$

$E(X) = \underset{\underset{1}{\parallel}}{E(X_1)} + \underset{\frac{n}{n-1}}{E(X_2)} + \cdots + \underset{\underset{n}{\parallel}}{E(X_n)}$

$= \boxed{n\left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}\right)}$