$\{X_n\}$: sequence of indpt Bern(p) RVs

$$N = \min\left\{k : \sum_{i=1}^{k} X_i = n\right\}$$

Find probability distn of N & $\sum_{i=1}^{N-1}(1-X_i)$

Hint

Suppose $n = 6$

Success - Failure run

0010100101001010001

$$\left\{k : \sum_{i=1}^{k} X_i = 6\right\}$$

$$= \{15, 16, 17, 18\}$$

$\therefore N = 15$

Note that $X_{15} = 1$ (6-th success at 15-th trial)

So $\sum_{i=1}^{N}(1-X_i) = \sum_{i=1}^{N-1}(1-X_i) = 9$ (here for the ex)

Consider a geometric series

$$\sum_{j=0}^{\infty} a^j$$ we know it converges for $|a| < 1$

Now, choose $0 < a < 1$

So $$\sum_{0}^{\infty} a^j = \frac{1}{1-a} > 0$$

$$\Rightarrow \sum_{0}^{\infty} (1-a)a^j = 1$$

Re write $(1-a) = p$

& define $f(j) = p(1-p)^j I_j\{0, 1, \ldots\}$

Clearly f is a PMF. (see the formulation)

What are the feature of such type of probability Model?

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

$$\Rightarrow \frac{d}{dq}\left(\quad '' \quad \right) = \frac{1}{(1-q)^2}$$ [$\because$ the above series is absolutely conv $\forall\, q \in (0,1)$]

or $$\sum_{1}^{\infty} j q^{j-1} = \frac{1}{(1-q)^2}$$

$$\Rightarrow \sum_{1}^{\infty} j p q^j = \frac{pq}{p^2} = \frac{q}{p}.$$

$\therefore$ If X has PMF f

Then $$E(X) = \sum_{0}^{\infty} j p q^j = \sum_{1} j p q^j = \frac{q}{p}.$$

|||ly $$\frac{d^2}{dq^2} \sum_{0}^{\infty} q^j = \frac{2}{(1-q)^3}$$

$$\Rightarrow \sum_{2}^{\infty} j(j-1) q^{j-2} = \frac{2}{p^3}$$ $\left[\text{Ans}\left(\frac{p}{q^2}\right) \right]$

$$\Rightarrow \sum_{2}^{\infty} j(j-1) p q^j = \frac{2pq^2}{p^3} = \frac{2q^2}{p^2}$$

$\therefore E(X(X-1)) = \dfrac{2q^2}{p^2}$ Now find $V(X) = ?$ *

We denote the distn by Geometric(p)

$\therefore$ Mean of Geometric (p) $= \frac{q}{p}$

& Var $= \frac{q}{p^2}$

Find $E(X \mid X \neq 0) = ?$

$V(X \mid X \neq 0) = ?$

* If $Y = X \mid X \neq 0$, then he e

$Y = 1 + X$ (a.s.) Verify ☆

PMF of $X \mid X \neq 0$ is

$$g(j) = \frac{p q^j}{P(X \neq 0)} I_j \{1, 2, \ldots\}$$

$$= \frac{p q^j}{1 - P(X = 0)} I_j \{1, 2, \ldots\}$$

$$= p q^{j-1} I_j \{1, 2, \ldots\}$$ Now check ☆

If $Y = X \mid X \neq 0$

$P(X=i) = p q^{i} \; I_i \{0, 1, \dots\}$

$P(Y=i) = p q^{i-1} \; I_i \{1, 2, \dots\}$

see that $Y = 1 + X$

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

$V(Y) = V(X) = \frac{q}{p^2}$.

Y is also a geometric RV.

This another form. You can also find $E(Y)$ & $V(Y)$ directly

A real life example of X:

No. of dry wells an oil company will drill in a given tenure before getting a productive well.

So we can generate the probability model using a seqn of Bern(p) trials, where X = # failures preceding the 1st success

& Y = # trials reqd to produce 1st success.

$$P(X = j) = P(\underbrace{FFF \cdots F}_{j}S)$$

$$= q^j p \, , \quad j = 0, 1, \ldots$$

$q = 1 - p$. (due to independence)

||,ly $P(Y=i) = P(\underbrace{FF\cdots F}_{i-1}S)$

$= q^{i-1}p, \quad i=1,2,\ldots\ldots$

(due to independence)

$q = 1-p.$

Now, in general we can find the factorial moments using PGF.

$$P_X(t) = \sum_0^\infty t^j p q^j$$

$$= p\sum_0^\infty (qt)^j$$

$$= \frac{p}{1-qt}$$

(Don't bother about convergence as $P_X(t)$ converges absolutely for $|t| \le 1$)

$\Rightarrow P_X(1+t) =$ factorial moment gen. fn.

$$= p(1-q(1+t))^{-1} = ?$$

Find $E(X)$, ($\checkmark$ NOW)

$$= \left(1 - \frac{q}{p} t\right)^{-1}$$

$$= \sum_r \left(\frac{q}{p} t\right)^r$$

$$\Rightarrow E(X)_r = r! \left(\frac{q}{p}\right)^r.$$

Problem

If $X = 0, 1, 2, \dots$, a.s. $\ni$

$$P(X \geq i+j \mid X \geq i) = P(X \geq j),$$

$$i, j = 0, 1, \dots$$

then show that $X \sim$ Geometric.

Let $G(j) = P(X \geqslant j)$

Note that $G(0) = 1$

$\therefore \; P(X \geqslant i+j \mid X \geqslant i) = P(X \geqslant j), \quad i, j = 0, 1, 2, \ldots$

$\Rightarrow P(X \geqslant i+j) / P(X \geqslant i) = P(X \geqslant j)$

$\Rightarrow \; G(i+j) = G(i)\,G(j), \quad i, j = 0, 1, 2, \ldots\ldots$

$$P(X = j) = P(X \geqslant j) - P(X \geqslant j+1)$$
$$= G(j) - G(j+1)$$

$$G(j+1) = G(j)\,G(1)$$
$$= G(j-1)\{G(1)\}^2$$
$$= G(j-2)\{G(1)\}^3$$
$$\vdots$$
$$= G(0)\{G(1)\}^{j+1} = \{G(1)\}^{j+1}$$

$$\therefore P(X = j) = \{G(1)\}^j - \{G(1)\}^{j+1}$$
$$= (1 - G(1))\{G(1)\}^j = p(1-p)^j, \quad j = 0, 1, 2, \ldots$$

$G(1) = P(X \geqslant 1) = 1 - P(X=0) = 1 - p$, say.

Suppose $X_1, X_2, \ldots, X_n$ are independently distributed geometric (p) random variables. Find the prob distn of $S_n = \sum_1^n X_i$.

Soln

$$P_{S_n}(t) = E(e^{tS_n})$$

$$= [P_{X_1}(t)]^n \quad (\because X_i\text{'s are IID})$$

Now, $P_{X_1}(t) = \sum_0^\infty t^j p q^j$, $q = 1-p$

$$= p \sum_0^\infty (qt)^j = p \frac{1}{1-qt}$$

$$\therefore P_{S_n}(t) = p^n (1-qt)^{-n}$$

$$= p^n \sum_0^\infty \frac{(-n)_j}{j!} (qt)^j (-1)^j$$

$\Rightarrow P(S_n = j) =$ coeff of t^j

$$= \frac{(-n)_j}{j!} (-1)^j p^n q^j, \quad j = 0, 1, 2, \cdots$$

$$\frac{(-n)_j (-1)^j}{j!} = \frac{-n(-n-1)\cdots(-n-j+1)(-1)^j}{j!}$$

$$= \frac{(n-1+j)(n-1+j-1)\cdots(n-1+j-j+1)}{j!}$$

$$= \binom{n-1+j}{j}$$

$$\therefore P(S_n = j) = \binom{n-1+j}{j} p^n q^j I_j\{0,1,\ldots\}$$

Understanding S_n

Let us consider a sequence of Bern(p) trials.
Repeat the trials till occurrence of n successes.
Now define

X_1 = # failures preceding the 1-st success.

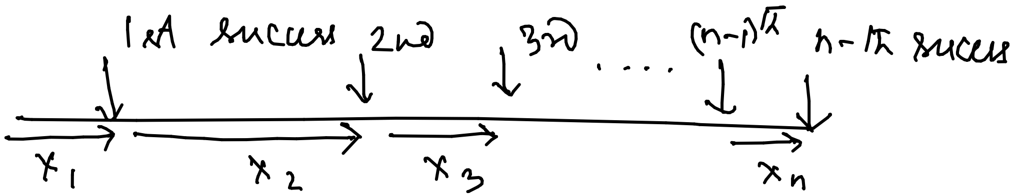
X_2 = # " preceding the 2-nd success after having the 1-st success

X_3 = # failures preceding the 3-rd success after having two successes & so on.

If $S_n = X_1 + X_2 + \cdots + X_n$,

What will be the interpretation of S_n in this case if X_i's are independently distn. ?

Clearly it is no. of failures preceding the n-th success & X_i's are all IID geometric (p) variables. So we can find the probability of $\{S_n (= X, \text{say}) = j\}$ directly.



Now $P(X=j) \quad (= P(S_n = j))$

$$= \binom{n-1+j}{j} P(\text{a sp case})$$

n-th S

$n-1 \to S$

$j \to F$

$$= \binom{n-1+j}{j} P(\underbrace{SSS\cdots S}_{n-1}\underbrace{FF\cdots F}_{j}\underbrace{\overset{\text{fixed}}{S}}_{1})$$

they can occur in any order

We have the multiplier $\binom{n-1+j}{j}$ as there are $\binom{n-1+j}{j}$ mutually exclusive and equally likely cases.

Now define Y = # trials required to produce n successes

Likewise decompose Y into geometric variables.

$$Y = Y_1 + Y_2 + \cdots + Y_n$$

Y_i = # trials reqd to produce a success after having $(i-1)$ successes, $i = 1, 2, \ldots, n$.

Y_i's are also IID with common PMF

$$f(j) = p\, q^{j-1}\, I_j\{1, 2, \ldots\}.$$

Here $Y = X + n$. a.s.

Obtain $P(Y = j)$ directly (HW)

$$P(Y=j) = P(n\text{-th success at the } j\text{-th trial})$$

$$= \binom{j-1}{n-1} P(\text{a specific case})$$

n-th S

n-1 successes

j-1 → trials (j-1) trial

$$= \binom{j-1}{n-1} P(\underbrace{SS\cdots S}_{n-1}\underbrace{FF\cdots F}_{j-n}\underbrace{S}_{1})$$

fixed.

$$= \binom{j-1}{n-1} p^n q^{j-n}, \quad j = n, n+1, \ldots$$

(To justify) HW

Now

Relationship between binomial and inverse binomial.

$P(X \le k)$ $\quad \{X \le k\} \equiv$ no of failures preceding n-th success is at the most k

$= P(Y \le n+k)$

$Y = X + n$

$\{Y \le n+k\} \equiv$ no of trials to have n successes is at the most $n+k$

If we deliberately perform $n+k$ trials there would be atleast n successes

Thus $P(X \leq k) = P(Y \leq n+k)$

$= P(U \geq n)$, $U \sim Bin(n+k, p)$

If we begin with $P(Y \leq k)$ then

$P(Y \leq k) = P(U \geq n)$, $U \sim Bin(k, p)$

Incomplete β' fn.

Known fact: $U \sim Bin(n, p)$

$P(U \leq k) = I_q(n-k, k+1)$,

where $n > k$, $k > -1$

Already derived in class (I guess)

Now just relate

$P(X \leq k) = P(U \geq n) = 1 - P(U \leq n-1)$

$= 1 - I_q(n+k-(n-1), (n-1)+1)$

$= 1 - I_q(k+1, n)$

$= I_p(n, k+1)$

$$P(X=j) = \binom{n-1+j}{j} p^n q^j \quad I_j\{0,1,\dots\}$$

$$P(X \le k) = \sum_0^k \binom{n-1+j}{j} p^n q^j$$

$$= \frac{1}{B(n,k+1)} \sum_0^k \binom{k}{j} p^n q^j \frac{(k-j)!\,(n-1+j)!}{(n+k)!}$$

$$= \frac{1}{B(n,k+1)} \sum_0^k \binom{k}{j} p^n q^j \times B(n+j, k-j+1)$$

to evaluate a must have $\binom{k}{j}$ ⊛ & then adjust accordingly

$$= \frac{1}{B(n,k+1)} \sum_0^k \binom{k}{j} p^n q^j \int_0^1 u^{n+j-1} (1-u)^{k-j}\, du$$

$$= \frac{1}{B(n,k+1)} \int_0^1 \left\{ \sum_0^k \binom{k}{j} (qu)^j (1-u)^{k-j} \right\} (pu)^{n-1} p\, du$$

$$= \frac{1}{B(n,k+1)} \int_0^1 (qu + 1-u)^k (pu)^{n-1} p\, du$$

$$= \frac{1}{B(n,k+1)} \int_0^1 (pu)^{n-1} (1-pu)^k p\, du$$

✓ done

replace pu by y

$E(X) = ?$ $\qquad V(X) = ?$

$$E(X) = \sum_{0}^{\infty} j \binom{n-1+j}{j} p^n q^j$$

$$= \sum_{0}^{\infty} j \binom{-n}{j} (-1)^j p^n q^j$$

$$= \sum_{1}^{\infty} \frac{-n}{j} \binom{-n-1}{j-1} (-1)^{j-1} (-1) \frac{p^{n+1}}{p} q^{j-1} q$$

$$= \frac{nq}{p} \sum_{1}^{\infty} \binom{-(n+1)}{j-1} p^{n+1} (-q)^{j-1}$$

$$= \frac{nq}{p} \underbrace{\sum_{0}^{\infty} \binom{-(n+1)}{i} p^{n+1} (-q)^{i}}_{1}$$

As $\displaystyle\sum_{0}^{\infty} \binom{n-1+j}{j} p^n q^j = \sum_{0}^{\infty} \binom{-n}{j} p^n (-q)^j = 1$

$\forall n \in \mathbb{N}$

Similarly $E[X(X-1)] = \sum_{0}^{\infty} j(j-1) \binom{-n}{j} p^{n} (-q)^{j}$

$$= \sum_{2}^{\infty} j(j-1) \frac{-n(-n-1)}{j(j-1)} \binom{-n-2}{j-2} \frac{p^{n+2}}{p^2} (-q)^{j-2} (-q)^{2}$$

$$= \frac{n(n+1)q^2}{p^2} \underbrace{\sum_{2}^{\infty} \binom{-\overline{n+2}}{j-2} p^{n+2} (-q)^{j-2}}_{1 \text{ (same argument)}}$$

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

$$\therefore V(X) = \frac{n(n+1)q^2}{p^2} + \frac{nq}{p} - \frac{n^2q^2}{p^2}$$

$$= \frac{nq^2}{p^2} + \frac{nq}{p} = \frac{nq}{p}\left(\frac{q}{p}+1\right)$$

$$= \frac{nq}{p^2}$$

Now $Y = X + n$ a.s.

$$\Rightarrow \quad E(Y) = n + E(X)$$
$$= n + \frac{nq}{p}$$
$$= \frac{n}{p}$$
$$V(Y) = V(X) = \frac{nq}{p^2}$$

HW

1. $E(Y) = ?$
2. $V(Y) = ?$

$\Big\}$ directly

3. $P(Y \le k) =$ in terms — incomplete β

4. $\frac{f(j+1)}{f(j)} = \frac{\alpha + \beta j}{j+1}$, directly

$j = 0, 1, \ldots$

$\beta \neq 1$

f has support $\{0, 1, \ldots\}$

Find PGF of the distn

Hence find the same for binomial, Poisson & Negative binomial.