Common discrete distributions

Usually, there are several equivalent ways to characterize a common r.v. X. (1) The story/ backgroud/ definition of X. (2) The explicit PMF/ CDF/ (PDF) of X. (3) Express X in terms of other r.v.s. The story shows the specialty of X and hints us the type of examples that we can use X to model. The PMF or CDF way is concise and allows us to do the computations.

Bernoulli r.v.. with parameter $p \in [0, 1]$. Denote $X \sim Bern(p)$.

(1) X is the outcome of a trial that succeeds with probability p and fails with probability $1 - p$.

(2) The PMF of X is $p(1) = P(X = 1) = p$ and $p(0) = P(X = 0) = 1 - p$.

Note $E[X] = p$ and $\mathrm{Var}(X) = p(1 - p)$.

Binomial r.v.. with parameter (n, p). Denote $X \sim Bin(n, p)$.

(1) X is the number of successes that occur in the n independent Bernoulli trials with parameter p.

(2) The PMF of X is $p(k) = P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k}$ for $k = 0, \ldots, n$.

(3) Let $X_1, \ldots, X_n$ be independent Bernoulli r.v.s with parameter p. Then $X = \sum_{k=1}^{n} X_k$.

Note $E[X] = np$ and $\mathrm{Var}(X) = np(1 - p)$.

Poisson r.v.. with parameter $\lambda > 0$. Denote $X \sim Poisson(\lambda)$.

(2) The PMF of X is $p(k) = P(X = k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$ for $k = 0, 1, 2, \ldots$

(1) Approximate binomial r.v. with parameter (n, p) when n large and p small such that np moderate. This story is not precise but flexible to model many examples.

Note $E[X] = \mathrm{Var}(X) = \lambda$.

Geometric r.v. & computing examples

Example 1 (Geometric r.v. with parameter p)**.** Denote $X \sim Geom(p)$.

(1) X is the number of independent Bernoulli trials with parameter $p \in (0, 1)$ such that first success occur.

(2) By independence, the PMF is $p(k) = (1 - p)^{k-1} p$ for $k = 1, 2, 3, \ldots$

Then we show $E[X] = \dfrac{1}{p}$ and $\mathrm{Var}(X) = \dfrac{1 - p}{p^2}$.

By definition, $E[X] = \sum_{k=1}^{\infty} k(1 - p)^{k-1} p$. Note that for $x \in (-1, 1)$, (by a result about the uniform convergence of power series) we have

$$\sum_{k=1}^{\infty} kx^{k-1} = \sum_{k=0}^{\infty} (x^k)' = (\sum_{k=0}^{\infty} x^k)' = (\frac{1}{1 - x})' = \frac{1}{(1 - x)^2}.$$

Then setting $x = 1-p$ leads to $E[X] = p\sum_{k=1}^{\infty} k(1-p)^{k-1} = p\dfrac{1}{[1-(1-p)]^2} = \dfrac{1}{p}$.

To obtain $\mathrm{Var}(X)$, it suffices to compute $E[X^2]$

$$\begin{aligned}
E[X^2] &= \sum_{k=1}^{\infty} k^2(1-p)^{k-1}p \\
&= \sum_{k=1}^{\infty}(k-1+1)^2(1-p)^{k-1}p \\
&= \sum_{k=1}^{\infty}(k-1)^2(1-p)^{k-1}p + \sum_{k=1}^{\infty} 2(k-1)(1-p)^{k-1}p + \sum_{k=1}^{\infty}(1-p)^{k-1}p \\
(\text{let } n = k-1) \quad &= (1-p)\sum_{n=1}^{\infty} n^2(1-p)^{n-1}p + 2(1-p)\sum_{k=1}^{\infty} n(1-p)^{n-1}p + 1 \\
&= (1-p)E[X^2] + 2(1-p)E[X] + 1
\end{aligned}$$

Then substitute $E[X] = \dfrac{1}{p}$ and solve the equation to get $E[X^2] = \dfrac{2-p}{p^2}$.

Hence $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \dfrac{1-p}{p^2}$.

Remark. Observe that we have use two different ways to compute $E[X]$ and $\mathrm{Var}(X)$ in Example 1 both of which can be recursively extended to deal with series like $\sum_{k=1}^{\infty} k^p x^k$ with $p \in \mathbb{N}$.

Example 2. Let $X \sim Bin(n,p)$. Prove

$$E\left[\frac{1}{1+X}\right] = \frac{1-(1-p)^{n+1}}{(n+1)p}$$

Proof. Recall the PMF of $Bin(n,p)$ and by the formula of $E[g(X)]$,

$$\begin{aligned}
E\left[\frac{1}{1+X}\right] &= \sum_{k=0}^{n} \frac{1}{k+1}\binom{n}{k} p^k(1-p)^{n-k} \\
&= \frac{1}{(n+1)p}\sum_{k=0}^{n} \frac{1}{k+1}\frac{(n+1)!}{k![(n+1)-(k+1)]!} \cdot p^{k+1}(1-p)^{[(n+1)-(k+1)]} \\
&= \frac{1}{(n+1)p}\sum_{k=0}^{n}\binom{n+1}{k+1} p^{k+1}(1-p)^{[(n+1)-(k+1)]} \\
(\text{let } j = k+1) \quad &= \frac{1}{(n+1)p}\left[\sum_{j=0}^{n}\binom{n+1}{j} p^j(1-p)^{[(n+1)-j]} - (1-p)^{n+1}\right] \\
(\text{ by Binomial Thm}) \quad &= \frac{1}{(n+1)p}[1-(1-p)^{n+1}].
\end{aligned}$$

□

Example 3. Let X be a r.v. with non-negative integral values. Prove

$$\sum_{k=0}^{\infty} kP(X \geq k) = \frac{1}{2}(E[X^2] + E[X]).$$

Proof.

$$\begin{aligned}\sum_{k=0}^{\infty} kP(X \geq k) &= \sum_{k=0}^{\infty}\sum_{i=k}^{\infty} kP(X = i)\\ &= \sum_{i=0}^{\infty}\sum_{k=0}^{i} kP(X = i)\\ &= \sum_{i=0}^{\infty}\frac{i(i+1)}{2}P(X = i)\\ &= \frac{1}{2}E[X^2] + \frac{1}{2}E[X],\end{aligned}$$

where in the second equality we have changed the order of summation (see e.g. [Tutorial 4, Example 2] for details). □

Remark. Recall for the r.v. X in Example 3 we also have the layer-cake $E[X] = \sum_{k=0}^{\infty} P(X \geq k)$ (see e.g. [Tutorial 4, Example 2]). Together we can express $E[X^2]$ in terms of $P(X \geq k)$ if $E[X] < \infty$, thus $\mathrm{Var}(X)$. This process can recursively continue to express $E[X^p], p \in \mathbb{N}$ in terms of $P(X \geq k)$.