Review of discrete random variables

I: General discrete distributions. For any discrete random variable Y, it has a pmf $p_Y(y) = P(Y = y)$ defined on its support R_Y:

1. For each y in R_Y, $p_Y(y) \geq 0$.

2. $\sum_{y \in R_Y} p_Y(y) = 1$

3. The expectation of Y is $E(Y) = \sum_{y \in R_Y} y p_Y(y)$.

4. The variance of Y is $V(Y) = E(Y^2) - \{E(Y)\}^2 = \sum_{y \in R_Y} y^2 p_Y(y) - \{E(Y)\}^2$.

5. Calculate a probability is a straightforward addition, for example, $P(Y \in B) = \sum_{y \in R_Y \cap B} p_Y(y)$.

6. The moment generating function of Y is

$$m_Y(t) = E(e^{tY}) = \sum_{y \in R_Y} e^{ty} p_Y(y).$$

   It can be used to generate moments of Y:

$$E(Y) = \left.\frac{dm_Y(t)}{dt}\right|_{t=0}, \quad E(Y^2) = \left.\frac{d^2 m_Y(t)}{dt^2}\right|_{t=0}, \quad V(Y) = \left.\frac{d^2 m_Y(t)}{dt^2}\right|_{t=0} - \left(\left.\frac{dm_Y(t)}{dt}\right|_{t=0}\right)^2.$$

II: Binomial distributions. It is used to model the number of "successes" out of n Bernoulli trials (where n is fixed). Notation: $Y \sim b(n, p)$, n is the number of trials, p is the success probability.

1. The support is $R_Y = \{0, 1, 2, \ldots, n\}$.

2. The pmf is $p_Y(y) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}$ for $y \in R_Y$. You can use TI-84 *binompdf*(n, p, y).

3. To calculate probability $P(Y \leq y)$ you can use IT-84 *binomcdf*(n, p, y).

4. The mgf of Y is $m_Y(t) = (1 - p + pe^t)^n$.

5. $E(Y) = np$ and $V(Y) = np(1-p)$.

III: Geometric distributions. It is used to model the number of trials to observe the first success. Notation: $Y \sim geom(p)$, p is the success probability.

1. The support is $R_Y = \{1, 2, \ldots, \infty\}$.

2. The pmf is $p_Y(y) = P(Y = y) = (1-p)^{y-1}p$ for $y \in R_Y$. This calculation is easy or TI-84 $geometpdf(p, y)$.

3. And $P(Y \leq y) = 1 - (1-p)^y$, again, an easy calculation or TI-84 $geometcdf(p, y)$.

4. The mgf of Y is
$$m_Y(t) = \frac{pe^t}{1-(1-p)e^t} \quad \text{for} \quad t < -\ln q.$$

5. $E(Y) = 1/p$ and $V(Y) = (1-p)/p^2$.

IV: Negative binomial distributions. It is used to model the number of trials to observe the rth success. Notation: $Y \sim nib(r, p)$, p is the success probability.

1. The support is $R_Y = \{r, r+1, r+2, \ldots, \infty\}$.

2. The pmf is
$$p_Y(y) = P(Y = y) = \binom{y-1}{r-1} p^r (1-p)^{y-r} \quad \text{for} \quad y \in R_Y.$$
This calculation can be done through $p \times binompdf(y-1, p, r-1)$.

3. $P(Y \leq y) = 1 - binomcdf(y, p, r-1)$.

4. The mgf of Y is
$$m_Y(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r \quad \text{for} \quad t < -\ln q.$$

5. $E(Y) = r/p$ and $V(Y) = r(1-p)/p^2$.

V: Hypergeometric distributions. We have a population of size N, out of N there are r objects of Class 1 and $N - r$ objects of Class 2. Randomly sample n objects out of N without replacement. Denote Y as the number of objects in the sample that belong to Class 1. Then Y has a hypergeometric distribution, written $Y \sim hyper(N, n, r)$.

1. The support is $R_Y = \{y : \max(0, n - N + r) \le y \le \min(n, r)\}$.
2. The pmf is
$$p_Y(y) = P(Y = y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \quad \text{for} \quad y \in R_Y.$$
3. And $P(Y \le y)$ is the summation of $p_Y(z)$ over $z \le y$.
4. $$E(Y) = n\left(\frac{r}{N}\right) \quad \text{and} \quad V(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right).$$

VI: Poisson distributions. It is used to model the number of occurrences in a certain unit of time or space. Notation: $Y \sim Poisson(\lambda)$, λ is the rate parameter.

1. The support is $R_Y = \{0, 1, 2, \ldots, \infty\}$.
2. The pmf is $p_Y(y) = P(Y = y) = \lambda^y e^{-\lambda}/y!$. You can use TI-84 *poissonpdf*(λ, y).
3. $P(Y \le y)$ can be calculated via *poissoncdf*(λ, y).
4. The mgf of Y is $m_Y(t) = \exp[\lambda(e^t - 1)]$.
5. $E(Y) = \lambda$ and $V(Y) = \lambda$.
6. Consistent in unit: If events or occurrences occur at a rate of λ per unit time or space, then the number of occurrences in an interval of length t units follows a Poisson distribution with rate λt.

VII: Approximation.

1. When N is large, you can use Binomial $b(n, p = r/N)$ to approximate Hypergeometric $hyper(N, n, r)$; i.e., if $Y \sim hyper(N, n, r)$ and N is large, then

$$P(Y = y) \approx binompdf(n, r/N, y), \quad P(Y \leq y) \approx binomcdf(n, r/N, y).$$

   Note that the calculation of $P(Y \leq y)$ can be very tedious if Y is Hypergeometrically distributed. This approximation can make such calculation much easier.

2. when n is large, one can use Poisson $poisson(\lambda = np)$ to approximate Binomial $b(n, p)$; i.e., if $Y \sim b(n, p)$ and n is large, then

$$P(Y = y) \approx poissonpdf(np, y), \quad P(Y \leq y) \approx poissoncdf(np, y).$$