Discrete Random Variables and Their Probability Distributions

Jeongyoun Ahn

KAIST

Random Variables

A real-valued function of each outcome in S.

Continuous RV: can take any real values from an interval.

Discrete: take finite or countable values

Notations

Uppercase letter, e.g., Y: A r.v.

lowercase letter, e.g., y: particular values that a r.v. may assume

$(Y = y) = \{(Y = y)\}$: the set of all points in S assigned the value y by the r.v. Y.

$P(Y = y)$, or $p(y)$: the sum of the probabilities of all sample points in S that are assigned the value y.

Probability Mass Function

$p(y) = P(Y = y)$

- $0 \leq p(y) \leq 1$, for all y

- $\sum_y p(y) = 1$.

Example 1) Of 8 candidates seeking three positions at a counseling center, 5 have degrees in social science and 3 do not. If 3 candidates are selected at random, find the probability distribution of Y, the number having social science degrees among the selected persons.

Expected Value

Expected value of Y:

$$E(Y) = \mu = \sum_{y} y p(y)$$

e.g., Binary r.v.

Expected value of a function of Y:

$$E(g(Y)) = \sum_{y} g(y) p(y)$$

Variance

$$\text{Var}(Y) = \sigma^2 = \text{E}[(Y - \mu)^2] \ \sigma = \sqrt{\text{Var}(Y)}$$

Example) A salesman of small business computer systems will contact three customers during a week. Each contact can result in either a sale, with probability .3, or no sale with probability .7. Assume that customer contacts are independent. $Y =$ number of computer systems sold during the week. Find the mean and the variance of Y.

Properties

a, b: constants

- $\mathrm{E}(ag(Y)+b) = a\mathrm{E}(g(Y)] + b$
- $\mathrm{Var}(ag(Y)+b) = a^2\mathrm{Var}[(g(Y)]$
- $\mathrm{E}[\sum_{i=1}^{k} g_i(Y)] = \sum_{i=1}^{k} \mathrm{E}[(g_i(Y)]$
- $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mu^2$

Binomial Experiment (Bernoulli trial)

Each trial results in one of two outcomes: success (S) or fail (F)

The probability of success on a single trial is equal to some value p and remains the same from trial to trial. The probability of a failure is $q = 1 - p$.

Binomial experiment consists of n identical, independent Bernoulli trials.

Binomial Distribution

Y: the number of successes observed during the n trials.

$$Y \sim B(n, p)$$

probability mass function:

$$p(y) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y = 0, 1, \ldots, n$$

$\mu = \mathrm{E}(Y) = np \ \sigma^2 = \mathrm{Var}(Y) = npq.$

Example

Suppose it is known that a new treatment is successful in curing a muscular pain in 60% of the cases. If it is tried on 15 patients, find the probability that

- At most 6 will be cured.
- The number cured will be no fewer than 6 and no more than 10.
- More than twelve will be cured.
- Exactly three will be cured.

Geometric Distribution

Y: the number of the trial on which the first success occurs under the binomial experiment.

$Y \sim Geo(p)$

$p(y) = q^{y-1}p, \quad y = 1, 2, \ldots$

$\mu = \mathrm{E}(Y) = 1/p$

$\sigma^2 = \mathrm{Var}(Y) = q/p^2$

Negative Binomial Distribution

Y: the number of the trial on which the rth success occurs under the binomial experiment,

$Y \sim NB(r, p)$

$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}, \quad y = r, r+1, r+2, \ldots$

$NB(1, p) = Geo(p)$

$\mu = \mathrm{E}(Y) = r/p$

$\sigma^2 = \mathrm{Var}(Y) = rq/p^2$

Hypergeometric Distribution

An urn contains N balls with r red and $N - r$ blue balls. Draw n balls and let Y be the number of red balls.

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

$y = 0, 1, 2, \ldots, n$, where $y \le r$ and $n - y \le N - r$.

Sampling with replacement (Binomial) vs. Sampling without replacement (Hypergeometric)

$\mu = \mathrm{E}(Y) = \frac{nr}{N}$

$\sigma^2 = \mathrm{Var}(Y) = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$

Poisson Distribution

Y: the number of rare events that occur in space, time, volume, and so on.

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, \ldots, \quad \lambda > 0$$

e.g.) the number of automobile accidents during a time period of one week

The binomial probability function approximates to the Poisson for large n, small p

$\mu = E(Y) = \lambda$

$\sigma^2 = V(Y) = \lambda$

Moments and Moments-Generating Funcition

k**th Moment**: $\mathrm{E}(Y^k)$

k**th Centered Moment**: $\mathrm{E}[(Y-\mu)^k]$

Moment-generating function of Y: $m_Y(t) = \mathrm{E}(e^{tY})$

$$\left.\frac{d^k m(t)}{dt^k}\right|_{t=0} = m^{(k)}(0) = \mathrm{E}(Y^k)$$

Tchebysheff's Inequality

Y: r.v. with mean μ and finite variance σ^2.

For any constant $k > 0$,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

or

$$P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$