STA255: Statistical Theory

Chapter 3: Discrete Random Variables and Their Probability Distributions (Part 2)

Summer 2017

Geometric Distribution

- Consider a sequence of independent Bernoulli trials.
- Each trial results in a 'success (S)' or 'failure (F)'. $P(S) = p$, $P(F) = 1 - p$.
- If we perform a **fixed** number n of these trials, then **the number of successes** is a random variable whose distribution is **Binomial(n, p).**
- Instead of fixing the number of trials n in advance, we wish to perform Bernoulli trials until the first success occurs.
- Then the number of successes becomes fixed at 1 and the number of trials is a random variable.

Geometric Distribution

- Examples:
  1. Number of tosses until heads show up for the first time.
  2. Number of driving tests until getting the driving licence.
  3. Number of times he will play until the first winning.
  4. Number times she tries her password until she find the right one.
  5. A search engine goes through a list of sites looking for a given key phrase. Number of sites visited until the key phrase is found.

Geometric Distribution

- Let Y = the number of the trial on which the first success occurs. We can derive the probability distribution of Y .
  (a) $Y \in \{1, 2, 3, 4, ...\}$
  (b) $p(y) = P(Y = y) = P(\{FF \dots FS\}) = (1 - p)^{y-1}p$. Here F is repeated y - 1 times.
- Such random variables are said to have the Geometric distribution with parameter p.
  Recall: p = probability of success.

Geometric Distribution

Definition (Geometric Distribution)

- The probability mass function of Geometric(p) is given by

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

- If $Y \sim Geometric(p)$, then:

$$E(Y) = \frac{1}{p}$$
$$V(Y) = \frac{1-p}{p^2}$$

Example

The probability that an applicant for drivers license passes the road test is 75%.

(a) What is the probability that an applicant passes the test on his fifth try?

Solution:

(b) What is the average and variance for the number of trials until he passes the road test?

Solution:

Negative Binomial Distribution

- We wish to perform Bernoulli trials until r successes occur.
- Then the number of successes becomes fixed at r and the number of trials is a random variable.
- Experiments of this kind are called negative binomial experiments.
- Examples:

1. Number of tosses until 5 heads show up.
2. Number times she tries her password until she finds the right one. [Geometric Distribution]

Negative Binomial Distribution

- Let Y = the number trials required to produce r successes in a negative binomial experiment.
- We can derive the probability distribution of Y as follows:
  (a) $y \in \{r, r+1, r+2, ...\}$
  (b) $P(Y = y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$.
- Special case: when r = 1, the distribution is called Geometric Distribution.

Mean and Variance

Definition (Mean and Variance)

- If $Y \sim NegBinom(r, p)$, then:

$$E(Y) = \frac{r}{p}$$

$$V(Y) = \frac{r(1-p)}{p^2}$$

Example

(a) Find the probability that a person flipping a coin gets the third head on the seventh flip.

Solution:

(b) Find the probability that a person flipping a coin gets the first head on the fourth flip.

Solution:

Hypergeometric Distribution

- Sampling without replacement. That is, trials are dependent.
- Hypergeometric experiment: we are interested in the probability of selecting y successes from the r items labeled successes and n - r failures from the N - r items labeled failures when a random sample of size n is selected from N items.
- Example: Consider a collection of N chips. (r chips are white and N - r chips are black). A collection of n chips are selected at random and without replacement. Find the probability that exactly y chips are white.

Solution

Solution: Y = number of white chips in the sample of n chips. By the multiplication principle we can write:

- Y in the previous example is said to have a hypergeometric distribution.
- $E(Y) = n(\frac{r}{N})$
- $V(Y) = n(\frac{r}{N})(1 - \frac{r}{N})\frac{N-n}{N-1}$

Example

A lot, consisting of 50 fuses, is inspected. If the lot contains 10 defective fuses what is the probability that in a sample of size 5

- there is no defective fuse.

  Solution: $Y =$ number of defective fuses in the sample.

- there are exactly 2 defective fuses.

  Solution:

Poisson Distribution

- The Poisson distribution models the number of occurrences of an event when there is a known average rate per unit time λ .
- Note: λ is a positive number. λ is the average number of events per unit of time.
- Examples:
    1. The number of visitors to a webserver per minute.
    2. The number of email messages received at the technical support center daily.
    3. The number of customers arriving at a service counter within one-hour period.
    4. The number of typographical errors in a book counted per page.
    5. The number of traffic accidents that occur on Ontario Highway 401 during a month.

Poisson Distribution

- The requirements for a Poisson distribution are that:
  (a) no two events can occur simultaneously (rare events),
  (b) events occur independently in different intervals, and
  (c) the expected number of events in each time interval remain constant.

Definition (Poisson Distribution)

- The probability mass function of Poisson(λ) is given by

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

- If $Y \sim Poisson(\lambda)$, then:

$$E(Y) = \lambda$$

$$V(Y) = \lambda$$

Example

Messages arrive at an electronic message center at random times, with an average of 9 messages per hour.

(a) What is the probability of receiving exactly five messages during the next hour?
Solution:

- R Output
  dpois(5,9)
  0.06072688

Example

(b) What is the probability that more than 10 messages will be received within the next two hours?

Solution:

- R Output

  ppois(10,18)

  0.03036626

  1-ppois(10,18)

  0.9696337

Relationship between Binomial Distribution and Poisson Distribution

Proposition (Binomial and Poisson)

If $Y \sim Bin(n, p = \lambda/n)$, then

$$\lim_{n \to \infty} \binom{n}{y} (\frac{\lambda}{n})^y (1 - \frac{\lambda}{n})^{n-y} = \frac{\lambda^y e^{-\lambda}}{y!}$$

The Moment Generating Function

Definition (Moments)

- The kth moment of a random variable Y taken about the origin is defined to be $E(Y^k), k = 1, 2, \ldots$
- $E(Y)$ is called the first moment about the origin.
- $E(Y^2)$ is called the second moment about the origin.
- $E(Y^k)$ is called the k moment about the origin.
- Definition: The kth moment of a random variable Y taken about its mean is defined to be $E[(Y-\mu)^k], k = 1, 2, \ldots$.
- $E[(Y-\mu)^k]$ is called the kth central moment about the mean.

The Moment Generating Function

Definition (Moment Generating Function)

The moment generating function (mgf) of a random variable Y, denoted by $m(t)$, is defined to be

$$m(t) = E(e^{tY}).$$

- We say that a moment-generating function for Y exists if there exists a positive constant b such that $m(t)$ is finite for $|t| \leq b$.

Moment Generating Function

- Note: $m(0) = E(e^{0\times Y}) = E(1) = 1$.
- One main use of the mgf is to find the moments of a random variable. That is $E(Y^k), k = 1, 2, \ldots$
- Note that,

$$m^{(1)}(t) = \frac{dm(t)}{dt} = \frac{d}{dt}E(e^{tY}) = E(\frac{d}{dt}e^{tY}) = E(Ye^{tY}).$$

- Set $t = 0$,we get $m^{(1)}(0) = E(Y)$.
- Similarly, $m^{(2)}(t) = E(Y^2e^{tY})$. Thus, $m^{(2)}(0) = E(Y^2)$.
- In general

Theorem

$$m^{(k)}(0) = E(Y^k).$$

where $m^{(k)}(0)$ is the kth derivative of $m(t)$ when $t = 0$.

Example

If $Y \sim Poisson(\lambda)$, find the moment generating function and use it to find the mean of this distribution.

Solution:

Example

Example: #3.155

Let $m(t) = (1/6)e^t + (2/6)e^{2t} + (3/6)e^{3t}$. Find the following:

(a) $E(Y)$

(b) $V(Y)$

(c) The distribution of Y .

Solution:

Example

Chebyshev's Inequality

- If we only know the mean and the variance for a probability distribution, then Chebyshev's inequality gives bounds (lower bound $1 - 1/k^2$ and upper bounds $1/k^2$) about probabilities for certain intervals about the mean.

Theorem (Chebyshev's Inequality)

Let Y be a random variable with mean μ and variance σ^2. Then for any positive k,

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

That is, at least $1 - 1/k^2$ of the distribution's values are within k standard deviations of the mean.

It follows that: $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$.

Example

The daily production of electric motors at a certain factory averaged 120 with a standard deviation of 10.

(a) What can be said about the fraction of days on which the production level falls between 100 and 140?

Solution:

Example

(b) Find the shortest interval certain to contain at least 90% of the daily production levels.

Solution:

Example

A toll bridge charges \$1.00 for passenger cars and \$2.50 for other vehicles. Suppose that during daytime hours, 60% of all vehicles are passenger cars. If 25 vehicles cross the bridge during a particular daytime period, what is the resulting expected toll revenue?

Solution:

Example

For a particular insurance policy the number of claims by a policy holder in 5 years is Poisson distributed. If the filing of one claim is four times as likely as the filing of two claims, find the expected number of claims.

Solution: