2.3 Moment Generating Function

Theorem 2.3.11

Let $F_X(x)$ and $F_Y(y)$ be two cdfs all of whose moments exist.

a. If X and Y have bounded supports, then $F_X(u) = F_Y(u)$ for all u if and only if $EX^r = EY^r$ for all integers $r = 0, 1, 2, \ldots$.

b. If the moment generating functions exist and $M_X(t) = M_Y(t)$ for all t in some neighborhood of 0, then $F_X(u) = F_Y(u)$ for all u.

Theorem 2.3.12 (Convergence of mgfs)

Suppose $\{X_i, i = 1, 2, \ldots\}$ is a sequence of random variables, each with mgf $M_{X_i}(t)$. Furthermore, suppose that

$$\lim_{i\to\infty} M_{X_i}(t) = M_X(t), \quad \text{for all } t \text{ in a neighborhood of } 0,$$

and $M_X(t)$ is an mgf.Then there is a unique cdf F_X whose moments are determined by $M_X(t)$ and, for all x where $F_X(x)$ is continuous, we have

$$\lim_{i\to\infty} F_{X_i}(x) = F_X(x).$$

That is, convergence, for $|t| < h$, of mgfs to an mgf implies convergence of cdfs.

Moment Generating Functions

- Binomial(n, p):

- Gamma(α, β):

- Normal(μ, σ):

- Uniform(a, b):

3.2 Discrete Distribution

(1) Discrete Uniform Distribution

A random variable X has a discrete uniform $(1, N)$ distribution if

$$P(X = x|N) = \frac{1}{N}, \quad x = 1, 2, \ldots, N,$$

where N is a specified integer. This distribution puts equal mass on each of the outcomes $1, 2, \ldots, N$.

We then have

$$EX = \sum_{x=1}^{N} xP(X = x|N) = \sum_{x=1}^{N} x\frac{1}{N} = \frac{N+1}{2},$$

and

$$EX^2 = \sum_{x=1}^{N} x^2\frac{1}{N} = \frac{(N+1)(2N+1)}{6},$$

and so,

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{(N+1)(2N+1)}{6} - (\frac{N+1}{2})^2 = \frac{(N+1)(N-1)}{12}.$$

(2) Hypogeometric Distriution

Suppose we have a large urn filled with N balls that are identical in every way except that M are red and $N-M$ are green. We reach in, blindfolded, and select K balls at random. Let X denote the number of red balls in a sample of size K, then X has a *hypergeometric distribution* given by

$$P(X = x|N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}, \quad x = 0, 1, \ldots, K.$$

The requirements for the range of X is

$$M \geq x \quad \text{and} \quad N - M \geq K - x,$$

which can be combined as

$$M - (N - K) \leq x \leq M.$$

The mean of this distribution is

$$EX = \sum_{x=0}^{K} x\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} = \sum_{x=1}^{K} x\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

Using the identities,

$$x\binom{M}{x} = M\binom{M-1}{x-1}$$
$$\binom{N}{K} = \frac{N}{K}\binom{N-1}{K-1},$$

we obtain

$$EX = \frac{KM}{N}\sum_{x=1}^{K}\frac{\binom{M-1}{x-1}\binom{N-M}{K-x}}{\binom{N-1}{K-1}} = \frac{KM}{N}.$$

A similar, but more lengthy, calculation will establish that

$$\text{Var}(X) = \frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right).$$

Example 3.2.1(Acceptance Sampling)

The hypergeometric distribution has application in acceptance sampling. Suppose a retailer buys goods in lots and each item can be either acceptable or defective. Let $N = \#$ of items in a lot, and $M = \#$ of defectives in a lot. Then we can calculate the probability that a sample of size K contains x defectives. To be specific, suppose that a lot of 25 machine parts is delivered, where a part is considered acceptable only if it passes tolerance. We sample 10 parts and find that none are defective (all are within tolerance). What is the probability of this event if there are 6 defectives in the lot of 25?

Applying the hypergeometric distribution with $N = 25$, $M = 6$, and $k = 10$, we have

$$P(X = 0) = \frac{\binom{6}{0}\binom{19}{10}}{\binom{25}{10}} = 0.028,$$

showing that our observed event is quite unlikely if there 6 defectives in the lot.

Example 3.2.A (Capture–Recapture Model)

- A comprehensive evaluation of capture-recapture models for estimating software defect content. Briand, L.C.; El Emam, K.; Freimut, B.G.;Laitenberger, O.; *Software Engineering, IEEE Transactions on , June 2000*

- Evaluating capture-recapture models with two inspectors. El Emam, K.; Laitenberger, O.; *Software Engineering, IEEE Transactions on , Sept. 2001*

- Fair bandwidth allocation for responsive and unresponsive flows using a capture-recapture model. Ming-Kit Chan; Hamdi, M.; *Communication Technology Proceedings, 2003. ICCT 2003. April 2003*

- An empirical study of experience-based software defect content estimation methods. Petersson, H.; Wohlin, C.; *Software Reliability Engineering, 1999. Proceedings. 10th International Symposium on. Nov. 1999*

- An active queue management scheme based on a capture-recapture model. Ming-Kit Chan; Hamdi, M.; *Selected Areas in Communications, IEEE Journal on , May 2003*

- Assessing software designs using capture-recapture methods. Vander Wiel, S.A.; Votta, L.G.; *Software Engineering, IEEE Transactions on , Nov. 1993*