Order Statistics

- The *order statistics* of a set of random variables $X_1, X_2, \dots, X_n$ are the same random variables arranged in increasing order.

- Denote by
  $X_{(1)}$ = smallest of $X_1, X_2, \dots, X_n$
  $X_{(2)}$ = 2nd smallest of $X_1, X_2, \dots, X_n$
  $\vdots$
  $X_{(n)}$ = largest of $X_1, X_2, \dots, X_n$

- Note, even if X_i's are independent, $X_{(i)}$'s can not be independent since $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$

- Distribution of X_i's and $X_{(i)}$'s are NOT the same.

Distribution of the Largest order statistic $X_{(n)}$

- Suppose $X_1, X_2, \dots, X_n$ are i.i.d random variables with common distribution function $F_X(x)$ and common density function $f_X(x)$.
- The CDF of the largest order statistic, $X_{(n)}$, is given by

$$F_{X_{(n)}}(x) = P\left(X_{(n)} \leq x\right) =$$

- The density function of $X_{(n)}$ is then

$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) =$$

Example

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d Uniform(0,1) random variables. Find the density function of $X_{(n)}$.

Distribution of the Smallest order statistic $X_{(1)}$

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables with common distribution function $F_X(x)$ and common density function $f_X(x)$.

- The CDF of the smallest order statistic $X_{(1)}$ is given by

$$F_{X_{(1)}}(x) = P\left(X_{(1)} \le x\right) = 1 - P\left(X_{(1)} > x\right) =$$

- The density function of $X_{(1)}$ is then

$$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) =$$

Example

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d Uniform(0,1) random variables. Find the density function of $X_{(1)}$.

Distribution of the kth order statistic $X_{(k)}$

- Suppose X_1, X_2,…, X_n are i.i.d random variables with common distribution function $F_X(x)$ and common density function $f_X(x)$.

- The density function of $X_{(k)}$ is

$$f_{X_{(n)}}(x) = \frac{n!}{(k-1)!(n-k)!}\left(F_X(x)\right)^{k-1}\left(1-F_X(x)\right)^{n-k} f_X(x)$$

Example

- Suppose X_1, X_2,…, X_n are i.i.d Uniform(0,1) random variables. Find the density function of $X_{(k)}$.

Some facts about Power Series

- Consider the power series $\sum_{k=0}^{\infty} a_k t^k$ with non-negative coefficients a_k.
- If $\sum_{k=0}^{\infty} a_k t^k$ converges for any positive value of t, say for $t = r$, then it converges for all t in the interval $[-r, r]$ and thus defines a function of t on that interval.
- For any t in $(-r, r)$, this function is differentiable at t and the series $\sum_{k=0}^{\infty} k a_k t^{k-1}$ converges to the derivatives.
- Example:
  For $k = 0, 1, 2, \ldots$ and $-1 < x < 1$ we have that

$$(1-x)^{-k-1} = \sum_{m=0}^{\infty} \binom{k+m}{k} x^m$$

(differentiating geometric series).

Generating Functions

- For a sequence of real numbers $\{a_j\} = a_0, a_1, a_2, \ldots$, the generating function of $\{a_j\}$ is

$$A(t) = \sum_{j=0}^{\infty} a_j t^j$$

if this converges for $|t| < t_0$ for some $t_0 > 0$.

Probability Generating Functions

- Suppose X is a random variable taking the values 0, 1, 2, … (or a subset of the non-negative integers).
- Let $p_j = P(X = j)$, j = 0, 1, 2, …. This is in fact a sequence $p_0, p_1, p_2, \ldots$
- **Definition:** The *probability generating function* of X is

$$\pi_X(t) = p_0 + p_1 t + p_2 t^2 + \cdots = \sum_{j=0}^{\infty} p_j t^j$$

- Since $\left|p_j t^j\right| \le p_j$ if $|t| < 1$ and $\sum_{j=0}^{\infty} p_j = 1$ the pgf converges absolutely at least for $|t| < 1$.
- In general, $\pi_X(1) = p_0 + p_1 + p_2 + \ldots = 1$.
- The pgf of X is expressible as an expectation:

$$\pi_X(t) = \sum_{j=0}^{\infty} p_j t^j = E\left(t^X\right)$$

Examples

- $X \sim$ Binomial(n, p),

$$\pi_X(t) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} t^j = (pt+q)^n$$

converges for all real t.

- $X \sim$ Geometric(p),

$$\pi_X(t) = \sum_{j=1}^{\infty} pq^{j-1} t^j = \frac{pt}{1-qt}$$

converges for $|qt| < 1$ i.e. $|t| < \frac{1}{q} = \frac{1}{1-p}$

Note: in this case $p_j = pq^j$ for $j = 1, 2, \ldots$

PGF for sums of independent random variables

- If X, Y are independent and $Z = X+Y$ then,

$$\pi_Z(t) = E\left(t^Z\right) = E\left(t^{X+Y}\right) = E\left(t^X t^Y\right) = E\left(t^X\right)E\left(t^Y\right) = \pi_X(t)\pi_Y(t)$$

- **Example**

  Let $Y \sim$ Binomial(n, p). Then we can write $Y = X_1 + X_2 + \ldots + X_n$. Where X_i's are i.i.d Bernoulli(p). The pgf of X_i is

$$\pi_{X_i}(t) = t^0(1-p) + t^1 p = tp + q.$$

  The pgf of Y is then

$$\pi_Y(t) = E\left(t^{X_1+X_2+\cdots+X_n}\right) = E\left(t^{X_1}\right)E\left(t^{X_2}\right)\cdots E\left(t^{X_n}\right) = (tp+q)^n.$$

Use of PGF to find probabilities

- **Theorem**

  Let X be a discrete random variable, whose possible values are the nonnegative integers. Assume $\pi_X(t_0) < \infty$ for some $t_0 > 0$. Then

$$\pi_X(0) = P(X = 0),$$

$$\pi'_X(0) = P(X = 1),$$

$$\pi''_X(0) = 2P(X = 2),$$

  etc. In general,

$$\pi_X^{(k)}(0) = k!\,P(X = k),$$

  where $\pi_X^{(k)}$ is the k^{th} derivative of π_X with respect to t.

- **Proof:**

Example

- Suppose $X \sim$ Poisson(λ). The pgf of X is given by

$$\pi_X(t) = \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^j}{j!} t^j =$$

- Using this pgf we have that

Finding Moments from PGFs

- **Theorem**

  Let X be a discrete random variable, whose possible values are the nonnegative integers. If $\pi_X(t) < \infty$ for $|t| < t_0$ for some $t_0 > 1$. Then

$$\pi'_X(1) = E(X),$$

$$\pi''_X(1) = E(X(X-1)),$$

  etc. In general,

$$\pi_X^{(k)}(1) = E(X(X-1)(X-2)\cdots(X-K+1)),$$

  Where $\pi_X^{(k)}$ is the kth derivative of π_X with respect to t.

- Note: $E(X(X\text{-}1)\cdots(X\text{-}k+1))$ is called the kth *factorial moment* of X.
- **Proof:**

Example

- Suppose $X \sim$ Binomial(n, p). The pgf of X is

$$\pi_X(t) = (pt+q)^n.$$

Find the mean and the variance of X using its pgf.

Uniqueness Theorem for PGF

- Suppose X, Y have probability generating function π_X and π_Y respectively. Then $\pi_X(t) = \pi_Y(t)$ if and only if $P(X = k) = P(Y = k)$ for $k = 0,1,2,\ldots$

- **Proof:**

  Follow immediately from calculus theorem:

  If a function is expressible as a power series at $x=a$, then there is only one such series.

  A pgf is a power series about the origin which we know exists with radius of convergence of at least 1.

Moment Generating Functions

- The *moment generating function* of a random variable X is

$$m_X(t) = E\left(e^{tX}\right)$$

$m_X(t)$ exists if $m_X(t) < \infty$ for $|t| < t_0 > 0$

- If X is discrete

$$m_X(t) = \sum_x e^{tx} p_X(x).$$

- If X is continuous

$$m_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx.$$

- Note: $m_X(t) = \pi_X(e^t)$.

Examples

- $X \sim$ Exponential(λ). The mgf of X is

$$m_X(t) = E\left(e^{tX}\right) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx =$$

- $X \sim$ Uniform(0,1). The mgf of X is

$$m_X(t) = E\left(e^{tX}\right) = \int_0^1 e^{tx} dx =$$

Generating Moments from MGFs

- **Theorem**

  Let X be any random variable. If $m_X(t) < \infty$ for $|t| < t_0$ for some $t_0 > 0$. Then

$$m_X(0) = 1$$
$$m'_X(0) = E(X),$$
$$m''_X(0) = E(X^2),$$

etc. In general,

$$m_X^{(k)}(0) = E(X^k),$$

Where $m_X^{(k)}$ is the kth derivative of m_X with respect to t.

- **Proof:**

Example

- Suppose $X \sim$ Exponential(λ). Find the mean and variance of X using its moment generating function.

Example

- Suppose $X \sim N(0,1)$. Find the mean and variance of X using its moment generating function.

Example

- Suppose $X \sim$ Binomial(n, p). Find the mean and variance of X using its moment generating function.

Properties of Moment Generating Functions

- $m_X(0) = 1$.
- If $Y=a+bX$, $a,b \in R$ then the mgf of Y is given by

$$m_Y(t) = E\left(e^{tY}\right) = E\left(e^{at+btX}\right) = e^{at}E\left(e^{btX}\right) = e^{at}m_X(bt).$$

- If X, Y independent and $Z = X+Y$ then,

$$m_Z(t) = E\left(e^{tZ}\right) = E\left(e^{tX+tY}\right) = E\left(e^{tX}e^{tY}\right) = E\left(e^{tX}\right)E\left(e^{tY}\right) = m_X(t)m_Y(t)$$

Uniqueness Theorem

- If a moment generating function $m_X(t)$ exists for t in an open interval containing 0, it uniquely determines the probability distribution.

Example

- Find the mgf of $X \sim N(\mu,\sigma^2)$ using the mgf of the standard normal random variable.

- Suppose, $X_1 \sim N\left(\mu_1, \sigma_1^2\right)$, $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$ independent.

  Find the distribution of X_1+X_2 using mgf approach.