# MATH 3338 Probability

Wenjiang Fu: wfu2@central.uh.edu

Lecture 10 - MATH 3338
Ch 10 Generating Functions

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Outline

1. Generating Functions for Discrete Distributions

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Generating Functions for Discrete Distributions

So far, we have seen many popular distributions, and calculate the mean and variance of the distributions.

**Q**: Can we determine a distribution completely if we know the mean and variance?

Many known distributions are determined by the distribution parameters. Such as $\mu$ and $\sigma$ in normal, $\lambda$ in Poisson, and $p$ in binomial, etc.

The answer is No. Then what features determine the distribution completely?

**Moments of discrete distributions**

For a given discrete distribution, define the $k$-th moment of the distribution by

$$\mu_k = E(X^k) = \sum_{j=1}^{\infty} (x_j)^k p(X = x_j),$$

provided the sum converges.

# Generating Functions for Discrete Distributions

- By the above definition, we see that the first moment and the second moment below.

$$\mu_1 = \sum_j (x_j) p(x_j) = \mu$$

$$\mu_2 = \sum_j (x_j^2) p(x_j) = Var(x) + \mu^2$$

  Other moments are also important. Such as the third moment, it measures the lack of symmetry of the distribution and is called the Kurtosis.

$$\mu_3 = \sum_j (x_j)^3 p(x_j)$$

- Can all moments (assuming all of them are finite) completely determine a distribution?
  Answer: Yes (in our textbook) but more precisely, under certain conditions.
  Correct answer: No, with no more conditions.

UNIVERSITY of HOUSTON DEPARTMENT OF MATHEMATICS

# Generating Functions for Discrete Distributions

- Multiple research publications show that all moments together $k = 1, 2, ...$, (assuming all exist), still do not completely determine the distribution, but only the tail part.
- However, under a bit more mild condition, the moments determine the distributions completely.
- **Moment Generating Functions**
  Define a function $g(t)$ of a variable $t$ in a small interval around 0, say $(-\varepsilon, \varepsilon)$,

$$g(t) = E(e^{tX}) = E\left(\sum_{k=0}^{\infty} \frac{(tX)^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{\mu_k t^k}{k!}$$

Also

$$E(e^{tX}) = \sum_{k=0}^{\infty} e^{tx_k} p(x_k)$$

# Generating Functions for Discrete Distributions

- **All moments can be easily generated with the generating function**
  Take the $n$-the derivative and evaluate it at $t = 0$, we will have the $n$-th moment of the distribution, assuming it exists.

$$g'(0) = E(Xe^{tX})|_{t=0} = E(X) = \mu_1$$

$$g''(0) = E(X^2 e^{tX})|_{t=0} = E(X^2) = \mu_2$$

$$\frac{d^n}{dt^n} g(t) = E(X^n) = \mu_n$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Generating Functions for Discrete Distributions

- **Examples**
  **Uniform** Assuming $X \sim Unif(1, 2, 3, , , , n)$, with prob $p(j) = 1/n$, for $j = 1, ..., n$. Then

$$g(t) = \sum_{j=1}^{n} \frac{1}{n} e^{tj} = \frac{1}{n}(e^t + e^{2t} + ... + e^{nt}) = \frac{e^t(e^{nt} - 1)}{n(e^t - 1)}.$$

  Taking derivative and setting $t = 0$, we have

$$\mu_1 = g'(0) = \frac{1}{n}(1 + 2 + .. + n) = \frac{n+1}{2}$$

$$\mu_2 = g''(0) = \frac{1}{n}(1 + 4 + 9 + ... + n^2) = \frac{(n+1)(2n+1)}{6}$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Examples**
  **Binomial** Assuming $X \sim Bin(n, p)$ with prob $p(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, for $k = 0, ..., n$. Then

$$g(t) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k}$$

$$= (pe^t + q)^n$$

  Taking derivative and setting $t = 0$, we have

$$\mu_1 = g'(0) = n(pe^t + q)^{n-1} pe^t|_{t=0} = np$$

$$\mu_2 = g''(0) = n(n-1)p^2 + np$$

- **Examples**
  **Geometric Distribution** Assuming $X \sim Geo(p)$ i.e., $p_X(k) = q^{k-1}p$, $k = 1, ...,$

$$g(t) = \sum_{k=1}^{\infty} e^{tk}(1-p)^{k-1}p = \frac{pe^t}{1-(1-p)e^t}$$

Taking derivative and setting $t = 0$, we have

$$\mu_1 = g'(0) = \frac{pe^t}{(1-(1-p)e^t)^2}|_{t=0} = \frac{1}{p}$$

$$\mu_2 = g''(0) = \frac{pe^t + p(1-p)e^{2t}}{(1-(1-p)e^t)^3}|_{t=0} = \frac{2-p}{p^2}$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

# Generating Functions for Discrete Distributions

- **Examples**

  **Poisson Distribution** Assuming $X \sim Pois(\lambda)$ i.e., $p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $k = 0, ...,$

$$g(t) = \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

  Taking derivative and setting $t = 0$, we have

$$\mu_1 = g'(0) = \lambda$$

$$\mu_2 = g''(0) = \lambda^2 + \lambda$$

# Generating Functions for Continuous Distributions

- **Moments**
  If $X$ is a continuous RV, on probability Space $\omega$, with density function $f_X$. Define the $n$-th moment by

$$\mu_n = E(X^n) = \int_{-\infty}^{\infty} x^n f_X(x)dx,$$

  provided that the integral $\int |x^n| f_X(x)dx$ is finite.
  First moment $\mu_1 = \mu = E(X)$, the second moment $\mu_2 = E(X^2) = \sigma^2 + \mu^2$.
- **Moment generating function**
  We define the moment generating function $g(t)$ for RV $X$ by

$$g(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx$$

$$= \sum_{k=0}^{\infty} \frac{\mu_k t^k}{k!} = \sum_{k=0}^{\infty} \frac{E(X^k)t^k}{k!}$$

  Provided this series converges.

# Generating Functions for Continuous Distributions

- **Moments**
  Then we have
  $$\mu_n = g^{(n)}(0).$$
- **Examples** Uniform on [0,1].
  $f(x) = 1$. Then
  $$\mu_n = \int_0^1 x^n dx = \frac{1}{n+1},$$
  $$g(t) = \int_0^1 e^{tx} dx = \frac{1}{t} e^{tx}|_0^1 = [e^t - 1]/t$$
  Exponential with parameter $\lambda > 0$.
  The generating function
  $$g(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}$$
  $$\mu_n = \frac{n!}{\lambda^n}$$

# Generating Functions for Continuous Distributions

- **Examples** Normal $N(0,1)$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

. The Moments and moment generating function are

$$\mu_n = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$= \frac{(2m)!}{2^m m!}, \text{ if } n = 2m; \quad \text{or } 0, \text{ if } n = 2m+1.$$

$$g(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tx - x^2/2} dx$$

$$= e^{t^2/2}$$

- for Normal $N(\mu, \sigma^2)$, we have

$$g(t) = e^{t\mu + t^2\sigma^2/2}$$

- **On Moment Generating Functions**
  Suppose RV $X$ has a moment generating function $g(t)$. Let $Y = aX + b$ with constants $a$ and $b$. Then, $Y$ has a generating function $g_Y(t) = e^{bt}g(at)$.
  **Solution** The generating function of $Y$

$$g_Y(t) = E(e^{tY}) = E(e^{t(aX+b)}) = E(e^{aXt}e^{bt}) = e^{bT}E(e^{(at)X})$$

$$= e^{bT}g_X(at)$$

- **Moment Problem**
  **Theorem 10.4** If $X$ is a bounded RV, then the moment generating function $g_X(t)$ of $x$ determines the density function $f_X(x)$ uniquely. Application of this Theorem is to prove the Central Limit Theorem (CLT): the mean of the IID RVs follows approximately a normal distribution,

# Generating Functions for Continuous Distributions

- **Moment Generating Functions for the proof of CLM**
  **MGF for standard normal distr** From previous results, we have the MGF of $N(0,1)$

  $$g(t) = e^{t^2/2}$$

  Suppose $X_1, ..., X_n$ are indep trials process with finite mean $\mu$ and variance $\sigma^2$. Let $S_n = X_1 + ... + X_n$ be the sum. The standardized sum is $S_n^* = (S_n - n\mu)/\sqrt{n\sigma^2}$. If $\mu = 0$, and $\sigma = 1$, then $S_n^* = S_n/\sqrt{n}$. We need to prove that the mean $S_n^*$ has its MGF converging to the normal MGF.
  Suppose each $X_1, ..., X_n$ has MGF $g(t)$. Then $S_n$ has a MGF $[g(t)]^n$ because $X_1, ..., X_n$ are indep and have the same MGF $g(t)$. Then the standardized sum $S_n^*$ has a MGF $g^*(t) = [g(t/\sqrt{n})]^n$. We would like to prove that $g^*(t) \to e^{t^2/2}$. Then the limit of $S_n^*(t)$ follows the standard normal.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

- **Moment Generating Functions for the proof of CLM**
  **MGF for standard normal distr**
  Notice that we do not know function $g(t)$, but assume it is smooth enough, and has all derivatives, and can take Taylor expansion at $t = 0$. Let function $u(t) = log(g(t))$. Then it has all derivatives, especially we have

$$u(0) = log(g(0)) = log(1) = 0, \;\; u'(0) = \frac{1}{g(0)} g'(0) = 0.$$

$$u''(t) = \frac{g''(t)g(t) - [g'(t)]^2}{(g(t))^2}$$

$$u''(0) = g''(0) \cdot 1 - [g'(0)]^2 = \mu_2 - \mu_1^2 = \sigma^2 = 1.$$

$$u^*(t) = nlog(g(\frac{t}{\sqrt{n}})) = n[u(0) + u'(0)\frac{t}{\sqrt{n}} + \frac{1}{2}u''(0)\frac{t^2}{n} + o(1)]$$

$$\to t^2/2 \;\text{ as }\; n \to \infty$$