Math/Stat 341: Probability: Fall ‘21 (Williams)

Professor Steven J Miller: sjm1@williams.edu

Homepage: https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21

Lecture 25: 11-15-21: https://youtu.be/HgupUwwgQ7M (slides)

Lecture 26: 11/08/19: Generating Functions and Moment Generating Functions: https://youtu.be/GxohirsuMfM (didn't do standardization / change of basis)

Plan for the day: Lecture 25: November 15, 2021:

https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21/handouts/341Notes_Chap1.pdf

- Generating Functions
- Moment Generating Functions
- Characteristic Functions
- Change of Base Formula

Telescoping sum:

$\cancel{13} - 1$
$+ \cancel{17} - \cancel{13}$
$+ \cancel{88} - \cancel{17}$
$+ \cancel{1701} - \cancel{88}$
$+ 2460 - \cancel{1701}$

24600

General items.

- Find the path through the algebra....

Definition 19.2.1 (Generating Function) *Given a sequence $\{a_n\}_{n=0}^{\infty}$, we define its generating function by*

$$G_a(s) = \sum_{n=0}^{\infty} a_n s^n$$

for all s where the sum converges.

Warning! $f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$

$f^{(n)}(0) = 0$ Taylor Series is identically zero, but f is not identically zero!

$\sin(x)$ and $1701 f(x) + \sin x$: same Taylor at $x = 0$

Theorem 19.3.1 (Uniqueness of generating functions of sequences) *Let $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ be two sequences of numbers with generating functions $G_a(s)$ and $G_b(s)$ which converge for $|s| < \delta$. Then the two sequences are equal (i.e., $a_i = b_i$ for all i) if and only if $G_a(s) = G_b(s)$ for all $|s| < \delta$. We may recover the sequence from the generating function by differentiating: $a_n = \frac{1}{n!}\frac{d^n G_a(s)}{ds^n}\Big|_{s=0}$.*

Term by Term: $G_a(s) = a_0 + \sum_{n=1}^{\infty} a_n s^n = b_0 + \sum_{n=1}^{\infty} b_n s^n$

Take $s=0 \Rightarrow a_0 = b_0$, Then $s \sum_{n=1}^{\infty} a_n s^{n-1} = s \sum_{n=1}^{\infty} b_n s^{n-1}$

Remove s: $a_1 + \sum_{n=2}^{\infty} a_n s^{n-1} = b_1 + \sum_{n=2}^{\infty} b_n s^{n-1}$.

$$\frac{G_a(s) - a_0}{s} = \frac{G_b(s) - b_0}{s}$$

Take $s = 0 \Rightarrow a_1 = b_1$

Shamp∞ mathematics...

Definition 19.4.1 (Convolution of sequences) *If we have two sequences* $\{a_m\}_{m=0}^{\infty}$ *and* $\{b_n\}_{n=0}^{\infty}$*, we define their convolution to be the new sequence* $\{c_k\}_{k=0}^{\infty}$ *given by*

$$c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_{k-1} b_1 + a_k b_0 = \sum_{\ell=0}^{k} a_\ell b_{k-\ell}.$$

We frequently write this as $c = a * b$.

$$\sum_{l=0}^{\infty} a_l x^l \cdot \sum_{n=0}^{\infty} b_n x^n = \sum_{k=0}^{\infty} c_k(a,b)\, x^k$$

x^0: $a_0 x^0 \cdot b_0 x^0 \to c_0 = a_0 b_0$

x^1: $a_0 x^0 \cdot b_1 x^1 + a_1 x^1 \cdot b_0 x^0 \to c_1 = a_0 b_1 + a_1 b_0$

Shampoo mathematics

Lemma 19.4.2 *Let* $G_a(s)$ *be the generating function for* $\{a_m\}_{m=0}^{\infty}$ *and* $G_b(s)$ *the generating function for* $\{b_n\}_{n=0}^{\infty}$. *Then the generating function of* $c = a * b$ *is* $G_c(s) = G_a(s)G_b(s)$.

$a * b * c = G_a(s)\, G_b(s)\, G_c(s) = G_{a*b*c}(s)$

Grouping: $(a*b)*c$

$G_{(a*b)*c}(s) = G_{a*b}(s)\, G_c(s)$

$= G_a(s)\, G_b(s)\, G_c(s)$

$$c_n = \sum_{\ell=0}^{n} a_\ell a_{n-\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell}\binom{n}{n-\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell}^2 \quad \sum_{k=0}^{2n} c_k s^k = G_c(s) = G_a(s)^2 = (1+s)^n \cdot (1+s)^n = (1+s)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} s^k$$

$a_\ell = \binom{n}{\ell}$ $\quad c = a * a \quad G_c(s) = G_a(s)\, G_a(s)$

$G_a(s) = \sum_{\ell=0}^{n} \binom{n}{\ell} s^\ell \cdot 1^{n-\ell} = (1+s)^n$ (n fixed)

$G_c(s) = G_a(s)\, G_a(s) = (1+s)^n (1+s)^n = (1+s)^{2n}$

But $(1+s)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} s^k 1^{2n-k}$

take $k=n$ get coeff of s^n is just $\binom{2n}{n}$

Proved $\sum_{\ell=0}^{n} \binom{n}{\ell}^2 = \sum_{\ell=0}^{n} \binom{n}{\ell}\binom{n}{n-\ell} = \binom{2n}{n}$

Story Proof: n Williams, n Amherst

$\binom{2n}{n}$ form a group of n people $= \sum_{\ell=0}^{n} \binom{n}{\ell} \binom{n}{n-\ell}$

choose ℓ from Williams

must take $n-\ell$ from Amherst

Definition 19.4.3 (Probability generating function) *Let X be a discrete random variable taking on values in the integers. Let $G_X(s)$ be the generating function to $\{a_m\}_{m=-\infty}^{\infty}$ with $a_m = \text{Prob}(X = m)$. Then $G_X(s)$ is called the probability generating function. If X is only non-zero at the integers, a very useful way of computing $G_X(s)$ is to note that*

$$G_X(s) = \mathbb{E}[s^X] = \sum_{m=-\infty}^{\infty} s^m \text{Prob}(X = m).$$

More generally, if the probabilities are non-zero on an at most countable set $\{x_m\}$, then

$$G_X(s) = E[s^X] = \sum_m s^{x_m} \text{Prob}(X = x_m).$$

Poisson λ

$$\text{Prob}(X = m) = \frac{\lambda^m e^{-\lambda}}{m!}$$

$$E[s^X] = \sum_{m=0}^{\infty} s^m \frac{\lambda^m e^{-\lambda}}{m!} = e^{-\lambda} \sum_{m=0}^{\infty} \frac{(\lambda s)^m}{m!} = e^{-\lambda} e^{\lambda s}$$

$$= e^{-\lambda(1-s)} = e^{\lambda(s-1)}$$

converges $\forall s \in \mathbb{R}$

(even $s \in \mathbb{C}$)

Theorem 19.4.4 *Let $X_1, \ldots, X_n$ be* independent *discrete random variables taking on non-negative integer values, with corresponding probability generating functions $G_{X_1}(s), \ldots, G_{X_n}(s)$. Then*

$$G_{X_1+\cdots+X_n}(s) = G_{X_1}(s) \cdots G_{X_n}(s).$$

Proof: $G_{X_1+\cdots X_n}(s) = E[s^{X_1+\cdots X_n}]$

$= E[s^{X_1} \cdots s^{X_n}]$

$= E[s^{X_1}] \cdots E[s^{X_n}]$

$= G_{X_1}(s) \cdots G_{X_n}(s)$

Shorter proof: adding is convolving use Lemma 19.4.2

Lemma 19.4.2 *Let $G_a(s)$ be the generating function for $\{a_m\}_{m=0}^{\infty}$ and $G_b(s)$ the generating function for $\{b_n\}_{n=0}^{\infty}$. Then the generating function of $c = a * b$ is $G_c(s) = G_a(s)G_b(s)$.*

The density of the sum of independent discrete random variables is the convolution of their probabilities!

Review, did this before

Since study sums all the time, shows the value of convolutions....

Definition 19.5.1 (Probability generating function) *Let X be a continuous random variable with density f. Then*

$$G_X(s) = \int_{-\infty}^{\infty} s^x f(x)dx$$

is the probability generating function of X.

Replace Σ with $\int$

$$G_X(s) = \mathbb{E}[s^X]$$

Maybe the correct general defn of a generating function is thru expectation

Definition 19.5.2 (Convolution of functions) *The convolution of two functions* f_1 *and* f_2*, denoted* $f_1 * f_2$*, is*

$$(f_1 * f_2)(x) = \int_{-\infty}^{\infty} f_1(t) f_2(x-t)\,dt.$$

If the f_i*'s are densities then the integral converges.*

X_i has density f_i then $X_1 + X_2$ has density $f_1 * f_2$ (at least if indep!)

Theorem 19.5.3 (Sums of continuous random variables) *The probability density function of the sum of independent continuous random variables is the convolution of their probability density functions. In particular, if* $X_1, \ldots, X_n$ *have densities* $f_1, \ldots, f_n$*, then the density of* $X_1 + \cdots + X_n$ *is* $f_1 * f_2 * \cdots * f_n$*.*

Proved when $n=2$, general case by grouping

Theorem 19.5.4 (Commutativity of convolution) *The convolution of two sequences or functions is* **commutative***; in other words,* $a*b = b*a$ *or* $f_1 * f_2 = f_2 * f_1$*.*

Trivial if f_1, f_2 are densities as have an interpretation

Then $X_1 + X_2 = X_2 + X_1$

In general: do some algebra/calculus

Definition 19.6.1 (Moments) *Let X be a random variable with density f. Its k^{th}* ***moment***, *denoted μ'_k, is defined by*

$$\mu'_k := \sum_{m=0}^{\infty} x_m^k f(x_m)$$

if X is discrete, taking non-zero values only at the x_m's, and for continuous X by

$$\mu'_k := \int_{-\infty}^{\infty} x^k f(x)dx.$$

In both cases we denote this as $\mu'_k = \mathbb{E}[X^k]$. We define the k^{th} ***centered moment***, *μ_k, by $\mu_k := \mathbb{E}[(X - \mu'_1)^k]$. We frequently write μ for μ'_1 and σ^2 for μ_2.*

Note $\mu_1 = 0$, $\mu_1' = \mu$

Definition 19.6.2 (Moment generating function) *Let X be a random variable with density f. The moment generating function of X, denoted $M_X(t)$, is given by $M_X(t) = \mathbb{E}[e^{tX}]$. Explicitly, if X is discrete then*

$$M_X(t) = \sum_{m=-\infty}^{\infty} e^{tx_m} f(x_m),$$

while if X is continuous then

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx. \quad = \mathbb{E}[e^{tX}]$$

Note $M_X(t) = G_X(e^t)$, or equivalently $G_X(s) = M_X(\log s)$.

When see terminology such as this, need to justify the name....

$$G_X(s) = \mathbb{E}[s^X]$$

$$M_X(s) = \mathbb{E}[e^{tX}]$$

$$s \leftrightarrow e^t$$

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{t^n x^n}{n!} f(x)dx = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[\int_{-\infty}^{\infty} x^n f(x)dx\right]$$ if converges / if $\int \Sigma = \Sigma \int$

$$= \sum_{n=0}^{\infty} \frac{\mu_n'}{n!} t^n$$

Theorem 19.6.3 *Let X be a random variable with moments μ'_k.*

1. *We have*

$$M_X(t) \;=\; 1+\mu'_1 t+\frac{\mu'_2 t^2}{2!}+\frac{\mu'_3 t^3}{3!}+\cdots;$$

   in particular, $\mu'_k = d^k M_X(t)/dt^k\Big|_{t=0}$.

2. *Let α and β be constants. Then*

$$M_{\alpha X+\beta}(t) \;=\; e^{\beta t} M_X(\alpha t).$$

   Useful special cases are $M_{X+\beta}(t) = e^{\beta t}M_X(t)$ and $M_{\alpha X}(t) = M_X(\alpha t)$; when proving the central limit theorem, it's also useful to have $M_{(X+\beta)/\alpha}(t) = e^{\beta t/\alpha}M_X(t/\alpha)$.

3. *Let X_1 and X_2 be independent random variables with moment generating functions $M_{X_1}(t)$ and $M_{X_2}(t)$ which converge for $|t| < \delta$. Then*

$$M_{X_1+X_2}(t) \;=\; M_{X_1}(t)M_{X_2}(t).$$

   More generally, if $X_1,\dots,X_N$ are independent random variables with moment generating functions $M_{X_i}(t)$ which converge for $|t| < \delta$, then

$$M_{X_1+\cdots+X_N}(t) = M_{X_1}(t)M_{X_2}(t)\cdots M_{X_N}(t).$$

   If the random variables all have the same moment generating function $M_X(t)$, then the right hand side becomes $M_X(t)^N$.

Theorem 19.6.5 ***(Uniqueness of moment generating functions for discrete random variables.)*** *Let X and Y be discrete random variables taking on non-negative integer values (i.e., they're non-zero only in $\{0, 1, 2, \ldots\}$) with moment generating functions $M_X(t)$ and $M_Y(t)$, each of which converges for $|t| < \delta$. Then X and Y have the same distribution if and only if there is an $r > 0$ such that $M_X(t) = M_Y(t)$ for $|t| < r$.*

There exist distinct probability distributions which have the same moments. In other words, knowing all the moments doesn't always uniquely determine the probability distribution.

Example 19.6.6 *The standard examples given are the following two densities, defined for $x \geq 0$ by*

$$\begin{aligned} f_1(x) &= \frac{1}{\sqrt{2\pi x^2}}\, e^{-(\log^2 x)/2} \\ f_2(x) &= f_1(x)\left[1 + \sin(2\pi \log x)\right]. \end{aligned} \tag{19.2}$$

It's a nice calculation to show that these two densities have the same moments; they're clearly different (see Figure 19.1).

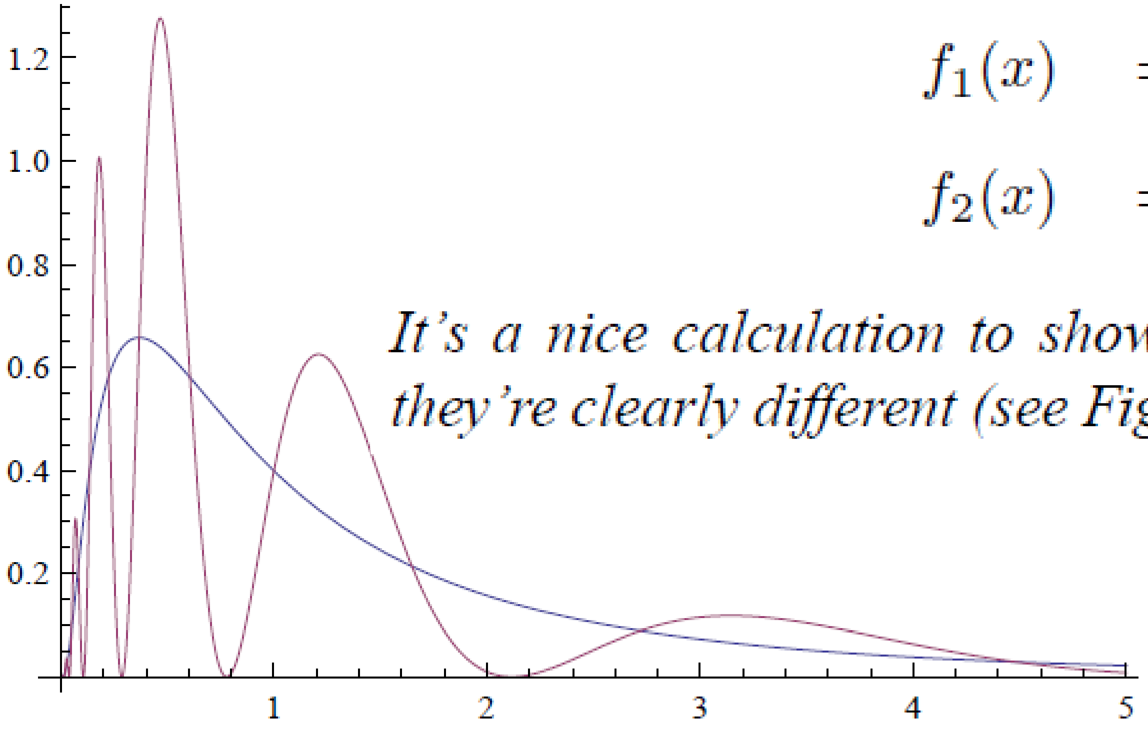

Figure 19.1: Plot of $f_1(x)$ and $f_2(x)$ from (19.2).

$$g(x) = \begin{cases} \exp(-1/x^2) & \text{if } x \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{19.3}$$

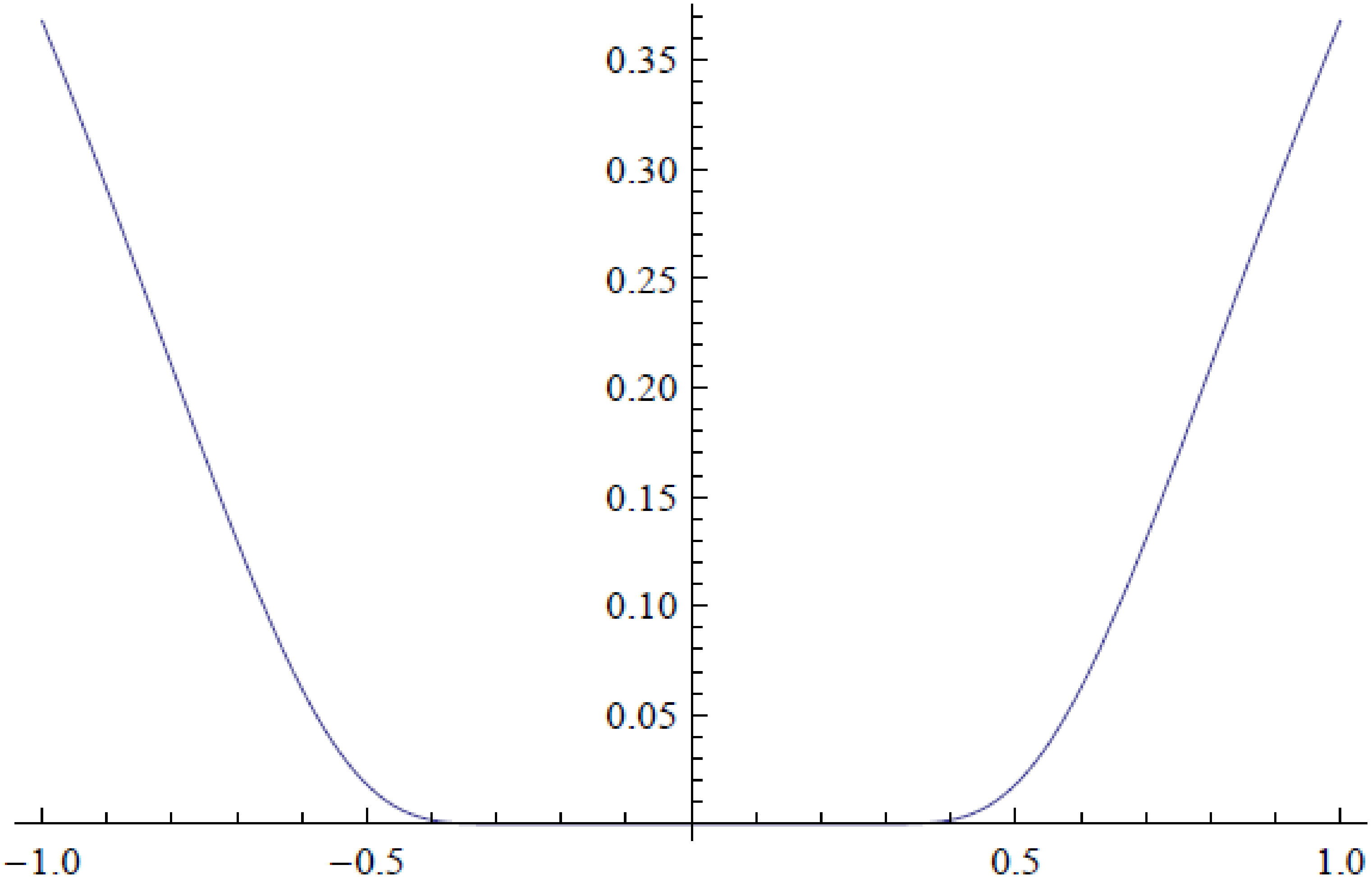

Figure 19.2: Plot of $g(x)$ from (19.3).

Poisson Generating Functions

$$M_X(t) = e^{\lambda(e^t - 1)} \qquad \mu = \frac{d}{dt} M_X(t)\Big|_{t=0} \qquad \mu_2' = \frac{d^2}{dt^2} M_X(t)\Big|_{t=0}$$

Change of Base Formula for Logarithms