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 23: 11-10-21: https://youtu.be/fY7teGfgxsY (slides)

Lecture 24: 11/04/19: Poisson Random Variables, Exponential Function, Stirling's Formula, Dyadic Decomposition, CLT to Stirling:
https://youtu.be/_-aGxDkNLHY

Plan for the day: Lecture 23: November 10, 2021:

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

- Theory of the Exponential Function
- Stirling's Formula (Dyadic Decomposition)
- Poisson Random Variables
- CLT to Stirling

General items.

- Power of differentiating identities
- Ability to use results in multiple ways

Definition 20.2.1 (Normal distribution) *A random variable X is normally distributed (or has the normal distribution, or is a Gaussian random variable) with mean μ and variance σ^2 if the density of X is*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

We often write $X \sim N(\mu, \sigma^2)$ to denote this. If $\mu = 0$ and $\sigma^2 = 1$, we say X has the standard normal distribution.

Theorem 20.2.2 (Central Limit Theorem (CLT)) *Let $X_1, \dots, X_N$ be independent, identically distributed random variables whose moment generating functions converge for $|t| < \delta$ for some $\delta > 0$ (this implies all the moments exist and are finite). Denote the mean by μ and the variance by σ^2, let*

$$\overline{X}_N = \frac{X_1 + \cdots + X_N}{N}$$

and set

$$Z_N = \frac{\overline{X}_N - \mu}{\sigma/\sqrt{N}}.$$

Then as $N \to \infty$, the distribution of Z_N converges to the standard normal (see Definition 20.2.1 for a statement).

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

$$\csc(x) = \frac{1}{\sin(x)}$$

$$e^x e^y = e^{x+y}$$

$$e^x e^y = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{x^n}{n!} \frac{y^m}{m!}$$

split by sum of powers

$$= \sum_{k=0}^{\infty} \sum_{l=0}^{k} \frac{x^l y^{k-l}}{l!\,(k-l)!} \frac{k!}{k!}$$

want $k = n+m$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \underbrace{\sum_{l=0}^{k} \binom{k}{l} x^l y^{k-l}}_{\text{Binomial Thm}}$$

$$= \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} = e^{x+y} \quad \blacksquare$$

The Gamma function. The Gamma function $\Gamma(s)$ is

$\Gamma(n+1) = n!$
$n = 0, 1, 2, \ldots$

$$\Gamma(s) = \int_0^\infty e^{-x} x^{s-1} dx, \quad \Re(s) > 0.$$

" $x^s \frac{dx}{x}$ "

Stirling's formula: As $n \to \infty$, we have

$$n! \approx n^n e^{-n} \sqrt{2\pi n};$$

by this we mean

$$\lim_{n\to\infty} \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = 1.$$

More precisely, we have the following series expansion:

$$n! = n^n e^{-n}\sqrt{2\pi n}\left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} - \cdots\right).$$

Crude upper/lower bounds for n!.

$$0 \le 1 \le n \le \quad n! \quad \le n^n$$

Thm: $\underbrace{1+2+3+\dots+n}_{} = \frac{n(n+1)}{2}$

$$n \le S(n) \le n^2$$

Lower bound: $\frac{n}{2}\cdot\frac{n}{2} = \frac{n^2}{4}$

have $\frac{n}{2} + \left(\frac{n}{2}+1\right) + \dots + n$: $\frac{n}{2}$ terms, each is $\frac{n}{2}$

$$\frac{1}{4}n^2 \le S(n) \le n^2$$

Note (n+1)!/n! = n+1; let's see what Stirling gives:

$$n! \approx n^n e^{-n} \sqrt{2\pi n}$$

$$(n+1)! \approx (n+1)^{n+1} e^{-(n+1)} \sqrt{2\pi(n+1)}$$

$$\frac{(n+1)!}{n!} \approx \frac{(n+1)^{n+1}}{n^n} \frac{e^{-n-1}}{e^{-n}} \frac{\sqrt{2\pi(n+1)}}{\sqrt{2\pi n}}$$

$$\approx (n+1) \underbrace{\left(1+\frac{1}{n}\right)^n e^{-1}}_{e \cdot 1/e} \sqrt{1+\frac{1}{n}} \simeq (n+1)$$

Integral Test and the Poor Mathematician's Stirling

factors are 1 and n

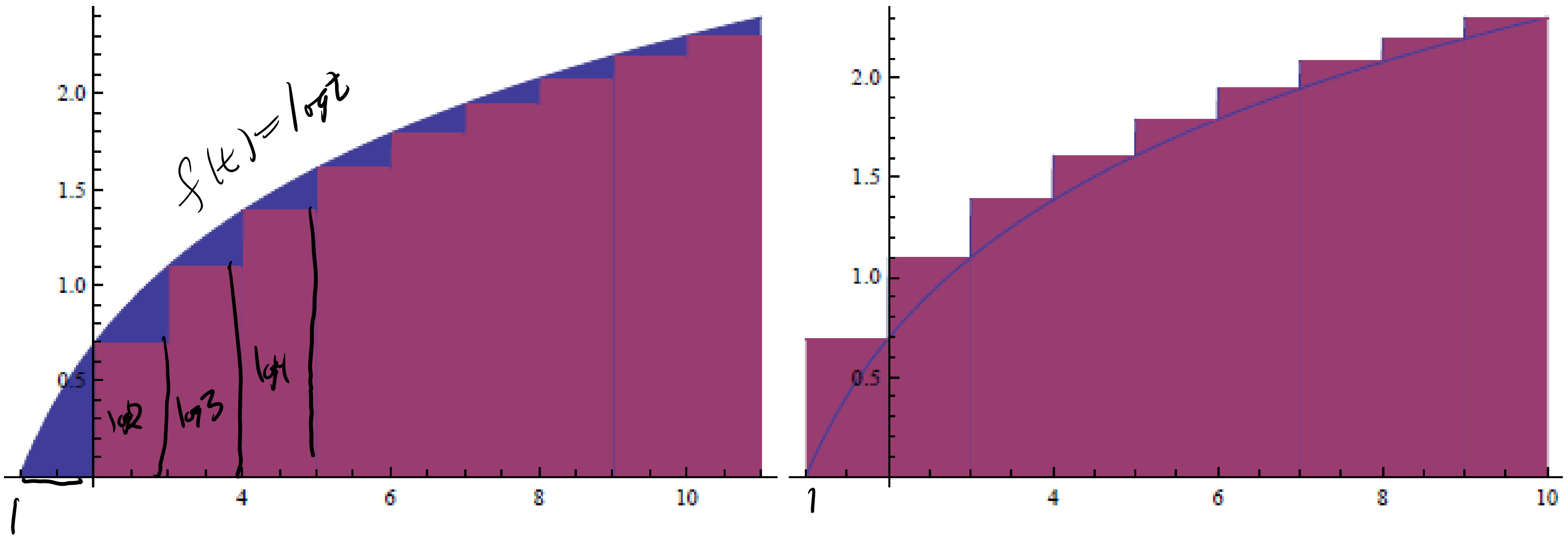
$GM(1,n) = \sqrt{1\cdot n} = \sqrt{n}$

Poor man's Stirling. Let $n \geq 3$ be a positive integer. Then

$$n^n e^{-n} \cdot e \;\leq\; n! \;\leq\; n^n e^{-n} \cdot en.$$

$$\log P \;=\; \log n! \;=\; \log 1 + \log 2 + \cdots + \log n \;=\; \sum_{k=1}^{n} \log k.$$

$$\int_1^n \log t dt \;\leq\; \sum_{k=1}^{n} \log k \;\leq\; \int_2^{n+1} \log t dt.$$



Lower and upper bound for $\log n!$ when $n = 10$.

Stirling's Formula: Lower bound from Integral Test:

Claim $(t\log t - t)' = \log t$

$$\begin{aligned}
(t\log t - t)\Big|_{t=1}^{n} \quad &\le \; \log n! \; \le \quad (t\log t - t)\Big|_{t=2}^{n+1} \\
n\log n - n + 1 \quad &\le \; \log n! \; \le \quad (n+1)\log(n+1) - (n+1) - (2\log 2 - 2).
\end{aligned}$$

We'll study the lower bound first. From

$$n\log n - n + 1 \;\le\; \log n!,$$

we find after exponentiating that

$$e^{n\log n - n + 1} \;=\; n^n e^{-n}\cdot e \;\le\; n!.$$

Euler–Maclaurin formula

From Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/Euler%E2%80%93Maclaurin_formula

If m and n are natural numbers and $f(x)$ is a real or complex valued continuous function for real numbers x in the interval $[m,n]$, then the integral

$$I = \int_m^n f(x)\,dx$$

can be approximated by the sum (or vice versa)

$$S = f(m+1) + \cdots + f(n-1) + f(n)$$

(see rectangle method). The Euler–Maclaurin formula provides expressions for the difference between the sum and the integral in terms of the higher derivatives $f^{(k)}(x)$ evaluated at the endpoints of the interval, that is to say $x = m$ and $x = n$.

Explicitly, for p a positive integer and a function $f(x)$ that is p times continuously differentiable on the interval $[m,n]$, we have

$$S - I = \sum_{k=1}^{p} \frac{B_k}{k!} \left(f^{(k-1)}(n) - f^{(k-1)}(m) \right) + R_p,$$

where B_k is the kth Bernoulli number (with $B_1 = \frac{1}{2}$) and R_p is an error term which depends on n, m, p, and f and is usually small for suitable values of p.

The formula is often written with the subscript taking only even values, since the odd Bernoulli numbers are zero except for B_1. In this case we have[1][2]

$$\sum_{i=m}^{n} f(i) = \int_m^n f(x)\,dx + \frac{f(n)+f(m)}{2} + \sum_{k=1}^{\lfloor \frac{p}{2} \rfloor} \frac{B_{2k}}{(2k)!} \left(f^{(2k-1)}(n) - f^{(2k-1)}(m) \right) + R_p,$$

or alternatively

$$\sum_{i=m+1}^{n} f(i) = \int_m^n f(x)\,dx + \frac{f(n)-f(m)}{2} + \sum_{k=1}^{\lfloor \frac{p}{2} \rfloor} \frac{B_{2k}}{(2k)!} \left(f^{(2k-1)}(n) - f^{(2k-1)}(m) \right) + R_p.$$

The Poisson distribution. Let $\lambda > 0$. Then X is a **Poisson random variable** with parameter λ if

$$\text{Prob}(X = n) = \begin{cases} \lambda^n e^{-\lambda}/n! & \text{if } n \in \{0, 1, 2, \dots\} \\ 0 & \text{otherwise.} \end{cases}$$

We write $X \sim \text{Pois}(\lambda)$. The mean and the variance are both λ.

Clearly non-neg

Does it sum to 1?

$$\sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \lambda^n/n! = e^{-\lambda} e^{\lambda} = 1$$

Save: $\sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} = 1 \implies \boxed{e^{\lambda} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}}$

Poisson random variable: $\mu_X = \lambda$

$$e^{\lambda} = \sum_{n=0}^{\infty} \lambda^n / n!$$

apply $\lambda \frac{d}{d\lambda}$

$$\lambda e^{\lambda} = \sum_{n=0}^{\infty} \lambda \frac{d}{d\lambda}\left(\frac{\lambda^n}{n!}\right)$$

$$= \sum_{n=0}^{\infty} \frac{n \lambda^n}{n!}$$

$$\lambda = \sum_{n=0}^{\infty} n \frac{\lambda^n e^{-\lambda}}{n!}$$

$$= E[X]$$

$$\boxed{E[X]} = \sum_{n=0}^{\infty} n \, \text{Prob}(X = n)$$

$$= \sum_{n=0}^{\infty} n \cdot \frac{\lambda^n e^{-\lambda}}{n!}$$

$$= \sum_{n=1}^{\infty} \frac{n}{n!} \lambda^n e^{-\lambda} \qquad (\lambda^n = \lambda^{n-1}\lambda)$$

$$= \sum_{n=1}^{\infty} \frac{\lambda \cdot \lambda^{n-1} e^{-\lambda}}{(n-1)!}$$

$$= \lambda \sum_{m=0}^{\infty} \frac{\lambda^m e^{-\lambda}}{m!}$$

$m = n-1$

$n: 1 \to \infty$

$m: 0 \to \infty$

$$= \lambda \cdot 1 = \lambda$$

Sums of Poisson random variables. The sum of n independent Poisson random variables with parameters $\lambda_1, \ldots, \lambda_n$ is a Poisson random variable with parameter $\lambda_1 + \cdots + \lambda_n$.

$$((X_1 + X_2) + X_3) + \cdots$$

$X_1 \sim Poiss(\lambda_1)$ $\quad$ $X_2 \sim Poiss(\lambda_2)$

$X = X_1 + X_2$

$$Prob(X = n) = \sum_{m=0}^{n} Prob(X_1 = m)\, Prob(X_2 = n-m)$$

discrete
convolution

$$= \sum_{m=0}^{n} \frac{\lambda_1^m e^{-\lambda_1}}{m!} \frac{\lambda_2^{n-m} e^{-\lambda_2}}{(n-m)!} \frac{n!}{n!}$$

$$= \frac{1}{n!} \underbrace{\sum_{m=0}^{n} \binom{n}{m} \lambda_1^m \lambda_2^{n-m}} e^{-(\lambda_1+\lambda_2)}$$

$$= \frac{(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}}{n!}$$

Definition 20.2.1 (Normal distribution) *A random variable X is normally distributed (or has the normal distribution, or is a Gaussian random variable) with mean μ and variance σ^2 if the density of X is*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

We often write $X \sim N(\mu, \sigma^2)$ to denote this. If $\mu = 0$ and $\sigma^2 = 1$, we say X has the standard normal distribution.

Theorem 20.2.2 (Central Limit Theorem (CLT)) *Let $X_1, \ldots, X_N$ be independent, identically distributed random variables whose moment generating functions converge for $|t| < \delta$ for some $\delta > 0$ (this implies all the moments exist and are finite). Denote the mean by μ and the variance by σ^2, let*

$$\overline{X}_N = \frac{X_1 + \cdots + X_N}{N}$$

and set

$$Z_N = \frac{\overline{X}_N - \mu}{\sigma/\sqrt{N}}.$$

Then as $N \to \infty$, the distribution of Z_N converges to the standard normal (see Definition 20.2.1 for a statement).

1. X has a Poisson distribution with parameter λ means

$$\text{Prob}(X = n) = \begin{cases} \frac{\lambda^n e^{-\lambda}}{n!} & \text{if } n \geq 0 \text{ is an integer} \\ 0 & \text{otherwise.} \end{cases}$$

2. If $X_1, \ldots, X_N$ are independent, identically distributed random variables with mean μ, variance σ^2 and a little more (such as the third moment is finite, or the moment generating function exists), then $X_1 + \cdots + X_N$ converges to being normally distributed with mean $n\mu$ and variance $n\sigma^2$.

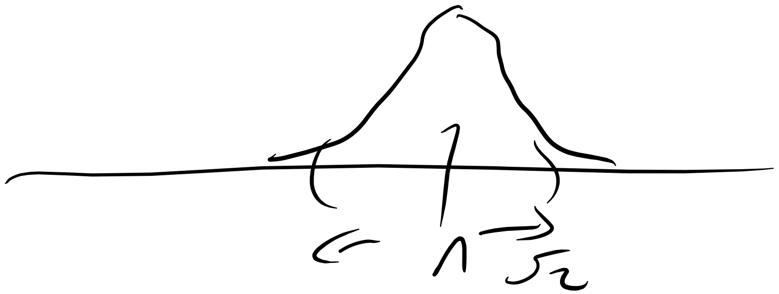
$X_k \sim \text{Poiss}(1)$

$E[X_k] = 1$

$\text{Var}(X_k) = 1$

$X = X_1 + \cdots + X_n \sim \text{Poiss}(n)$

$E[X] = n \qquad \text{Var}(X) = n$



X converges to $N(n, n) = N(\mu, \sigma^2)$

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

X approx $\mathcal{N}(n, n)$

$\text{Prob}\left(X \in [n-1/2,\ n+1/2]\right)$

$$\int_{n-1/2}^{n+1/2} \frac{1}{\sqrt{2\pi n}} e^{-(x-n)^2/2n}\, dx$$

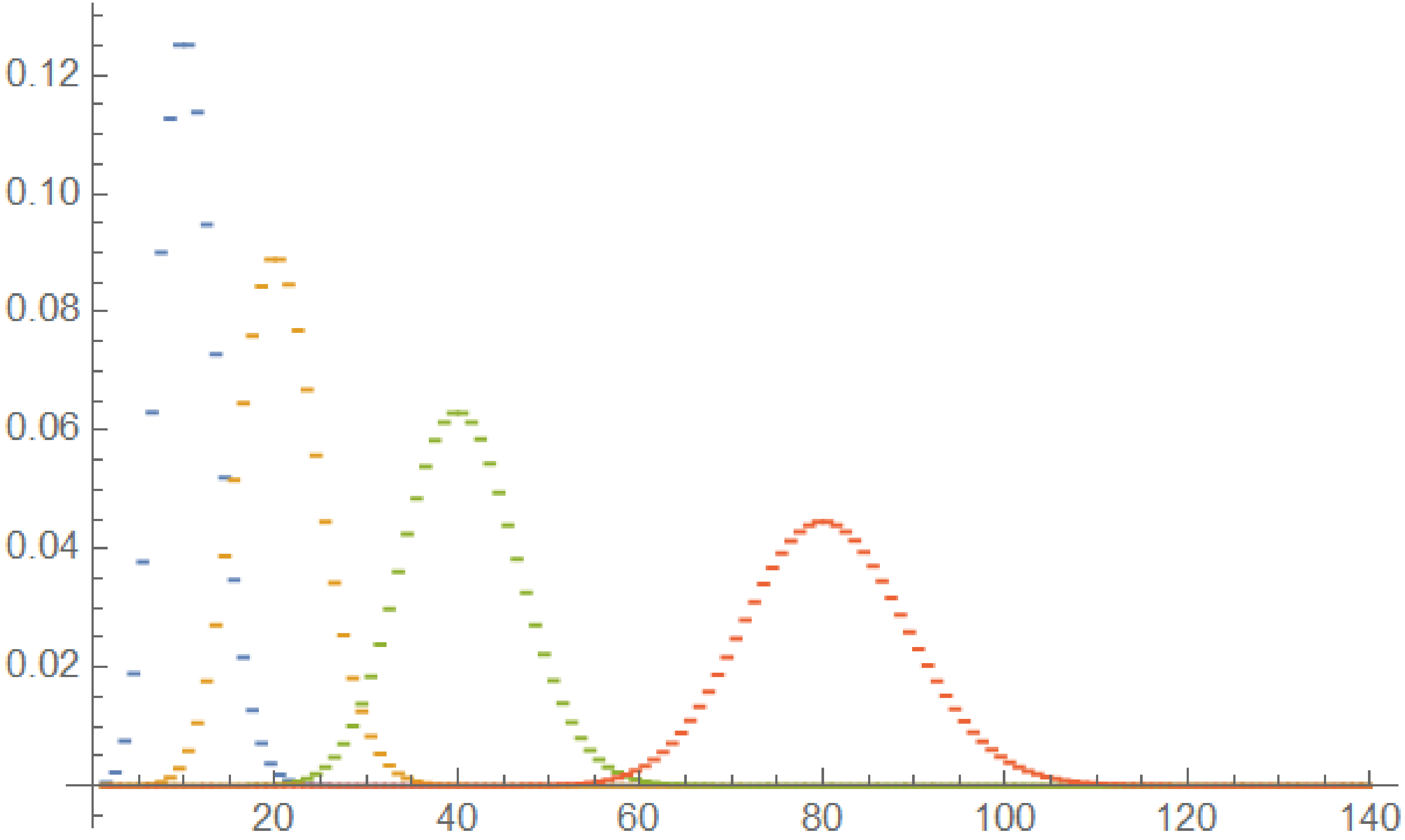
$e^{-1/8n} \approx 1$

$$\approx \int_{n-1/2}^{n+1/2} \frac{1}{\sqrt{2\pi n}}\, 1\, dx = \frac{1}{\sqrt{2\pi n}}$$

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

So $\displaystyle \frac{n^n e^{-n}}{n!} \approx \frac{1}{\sqrt{2\pi n}} \Rightarrow n! \approx n^n e^{-n} \sqrt{2\pi n}$

```
g[x_, lambda_] := Exp[-lambda] lambda^(Floor[x]) / (Floor[x]!)
Plot[{g[x, 10], g[x, 20], g[x, 40], g[x, 80]}, {x, 1, 140}, PlotRange → All]
```



g[x_, lambda_] := Exp[-lambda] lambda^(Floor[x])/ (Floor[x]!)
Plot[{g[x, 10], g[x, 20], g[x, 40], g[x, 80]}, {x, 1, 140}, PlotRange -> All]

$$\lim_{x \to 0} \frac{\sin x}{x} \underset{\text{L'Hopital}}{=} \lim_{x \to 0} \frac{\cos x}{1} = 1$$

Careful: need $(\sin x)' = \cos x$

$$\lim_{h \to 0} \frac{\sin(0+h) - \sin(0)}{h} = \lim_{h \to 0} \frac{\sin h}{h} = \lim_{x \to 0} \frac{\sin x}{x}$$

$$\lim_{n \to \infty} \frac{\sin x}{n} = 6$$