GAUSSIAN LIMIT OF POISSON DIST.

In last lecture [L13.] we noticed that for large λ, the Poisson distrib. resembled a Gaussian. To explain this we need to look at

$$P_n = \frac{\lambda^n}{n!} e^{-\lambda}$$

near its maximum as a function of n; i.e., within a few standard deviations $\sqrt{\lambda}$ of the mean, λ.

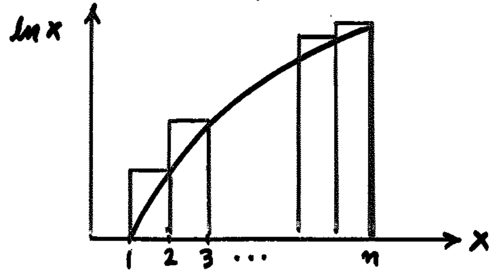
The awkward factor in P_n is the $n!$: we need an analytical approx. to it [so that we can differentiate it!], i.e. STIRLING'S APPROXIMATION.

STIRLING'S APPROXIMATION

Consider the integral

$$\begin{aligned} I &= \int_1^n \ln x \, dx \\ &= \int_1^n 1 \ln x \, dx \\ &= [x \ln x]_1^n - \int_1^n x \cdot \tfrac{1}{x} \, dx \quad \text{[by parts]} \\ &= n \ln n - n + 1 . \end{aligned}$$

Approximate I by sum of rectangles:



Tall rectangles:

$$I < \ln 2 + \ln 3 + \cdots + \ln n = \ln n!$$

Short rectangles: [see animation]

$$I > \ln 1 + \ln 2 + \ldots + \ln(n-1) \qquad \text{(A)}$$
$$= \ln n! - \ln n$$

[STIRLING'S APPROX.]

Hence

$$\ln(n!) > I = n\ln n - n + 1$$

and

$$\ln(n!) < I + \ln n \quad \text{[from (A)]}$$
$$= n\ln n - n + 1 + \ln n$$

gives lower and upper bounds on $\ln(n!)$. We can get a reasonable approx. for $\ln(n!)$ by averaging the upper and lower bounds:

$$\ln(n!) \simeq n\ln n - n + 1 + \ln\sqrt{n}$$

A more precise argument corrects the $+1$ to $\ln\sqrt{2\pi}$ [$\simeq 0.919$], giving

STIRLING'S APPROXIMATION:

$$\boxed{\ln(n!) \simeq n\ln n - n + \ln\sqrt{2\pi n}}$$

DON'T LEARN

It's a good approx, even for small n, and improves with increasing n.

E.g. $\ln 1! \simeq 1\times\ln 1 - 1 + \ln\sqrt{2\pi}$

$$\Rightarrow 1! \simeq \frac{\sqrt{2\pi}}{e} = 0.922$$ (close to the exact value!)

LIMIT OF POISSON DISTRIBUTION:

Because of the form of Stirling's approx [and because we're expecting P_n to have Gaussian form] we work with $f(n) = \ln P_n$ instead of P_n directly:

$$f(n) = \ln\left\{\frac{\lambda^n}{n!}e^{-\lambda}\right\}$$
$$= n\ln\lambda - \lambda - \ln(n!)$$
$$\simeq n\ln\lambda - \lambda - \{n\ln n - n + \ln\sqrt{2\pi} + \tfrac{1}{2}\ln n\}$$

MAX. w.r.t. n: at $n = \hat{n}$, with $f'(\hat{n}) = 0$

$$f'(n) = \ln\lambda - \ln n - \frac{1}{2n}$$

For $\lambda \gg 1 \Rightarrow \hat{n} \simeq \lambda$ as a first approx, since the "correction" $-\frac{1}{2\hat{n}} \simeq -\frac{1}{2\lambda}$ is small compared with the other terms.

[Second approx.: $\ln\lambda - \ln\hat{n} - \frac{1}{2\lambda} = 0$

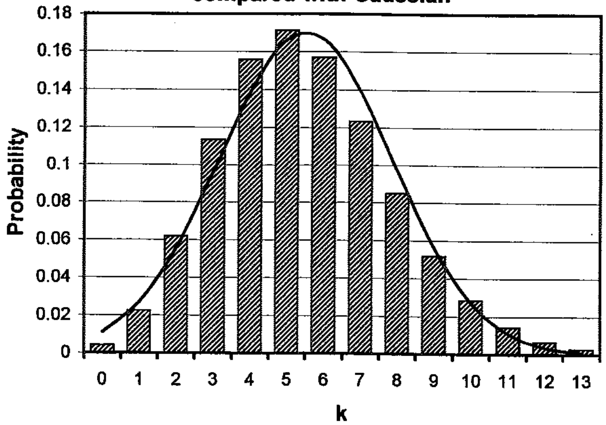
and note that $-\frac{1}{2\lambda} \simeq \ln\left(1 - \frac{1}{2\lambda}\right)$

giving $\ln\hat{n} \simeq \ln\lambda + \ln\left(1 - \frac{1}{2\lambda}\right)$

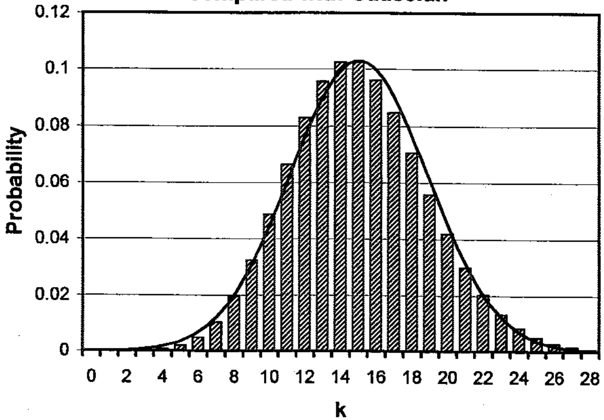
$$= \ln\left[\lambda\left(1 - \frac{1}{2\lambda}\right)\right] = \ln\left(\lambda - \tfrac{1}{2}\right)$$

$\Rightarrow \hat{n} \simeq (\lambda - \frac{1}{2})$. We <u>won't</u> use this.]

Poisson distribution, mean=5.5, compared with Gaussian



Poisson distribution, mean=15.0, compared with Gaussian



[LIMIT OF POISSON, CONT$^{\underline{D}}$]

TAYLOR EXPANSION ABOUT $n = \hat{n}$:

$$f(n) \simeq f(\hat{n}) + \frac{(n-\hat{n})}{1!}\underbrace{f'(\hat{n})}_{0} + \frac{(n-\hat{n})^2}{2!}f''(\hat{n}) + \dots$$

With "first approx" for $\hat{n}$,

$$f(\hat{n}) \simeq f(\lambda) = -\ln\sqrt{2\pi\lambda}$$

and

$$f''(\hat{n}) = \frac{d}{dn}\left\{\ln\lambda - \ln n - \frac{1}{2n}\right\}\Big|_{n=\lambda}$$

$$= \left\{-\frac{1}{n} + \frac{1}{2n^2}\right\}\Big|_{n=\lambda}$$

$$\simeq -\frac{1}{\lambda} \quad \left[\text{i.e., neglect } \frac{1}{2\lambda^2}\right]$$

Hence

$$f(n) \simeq -\ln\sqrt{2\pi\lambda} - \frac{(n-\lambda)^2}{2\lambda}$$

or

$$\boxed{P_n = e^{f(n)} \simeq \frac{e^{-(n-\lambda)^2/2\lambda}}{\sqrt{2\pi\lambda}}} \qquad (B)$$

GAUSSIAN LIMIT OF POISSON DISTRIBUTION

[LIMIT OF POISSON, CONT$^{\underline{D}}$]

RANGE OF VALIDITY:

For the Taylor series to be approximated well by the terms retained, we need the next correction $\frac{1}{3!}(n-\lambda)^3 f'''(\lambda)$ to be small compared with $\frac{1}{2}(n-\lambda)^2 \cdot \frac{1}{\lambda}$. But $f'''(\lambda) = \frac{1}{\lambda^2}$, so this requires

$|n-\lambda| \ll \lambda$. [Typically we have $|n-\lambda| \sim$ a few standard deviations $\sim \sqrt{\lambda} \ll \lambda$ for $\lambda \gg 1$, so no problem.]

[In addition, we effectively neglected $(n-\lambda) f'(\lambda) = (n-\lambda)\times\left(-\frac{1}{2\lambda}\right)$ when we used the 1st approx to $\hat{n}$. For this to be small compared with $(n-\lambda)^2/2\lambda$ we also need $|n-\lambda| \gg 1$. This is also fine for $|n-\lambda| \sim$ s.d. $\sim \sqrt{\lambda}$.]

A NOTE ON THE RESULT

We can think of the Gaussian limit of the Poisson distribution as a special case of the "central limit theorem".

Consider a random variable n, defined to be the sum of N Poisson-distributed random variables n_i, each with mean λ_i, which could be small:

$$n = \sum_{i=1}^{N} n_i, \quad \langle n_i \rangle = \lambda_i, \quad \mathrm{var}\, n_i = \sigma_i^2 = \lambda_i.$$

Now, the sum of Poisson random variables is also a Poisson random variable [see Lecture 13], with mean

$$\langle n \rangle = \sum_{i=1}^{N} \langle n_i \rangle = \sum_{i=1}^{N} \lambda_i \equiv \lambda$$

and $\mathrm{var}\, n = \lambda = \sum_{i=1}^{N} \sigma_i^2$ [since $\sigma_i^2 = \lambda_i$]

As N increases, λ becomes large, so result (B) tells us that $n = \sum n_i$ is Gaussian-distributed — even though the individual n_i's were not!! [The same general result holds, regardless of the distributions that the n_i's come from, provided σ_i^2 is finite and N is sufficiently large.]