M362K: February 28th, 2024.

Normal Approximation to the Binomial Dist'n [cont'd].

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{for } z \in \mathbb{R}$$

Standard Normal Density

Introduce: μ... mean parameter

σ... standard deviation parameter

Standard units		Raw measurements
$z = \frac{x-\mu}{\sigma}$	$\longleftrightarrow$	$x = \mu + \sigma \cdot z$

In general, the normal density is of the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for all } x \in \mathbb{R}$$

Example. Consider a normal distribution w/ mean 1 and standard deviation 2.
What is the probability of the interval $\boxed{[-0.5, 1.5]}$ under this distribution?

$\longrightarrow$:

$$\overset{c}{[-0.5}, \overset{d}{1.5]}$$

$$\frac{c-\mu}{\sigma} = \frac{-0.5-1}{2} = -\frac{3}{4} = -0.75$$

$$\frac{d-\mu}{\sigma} = \frac{1.5-1}{2} = \frac{1}{4} = 0.25$$

answer: $\Phi(0.25) - \Phi(-0.75) =$

$$= 0.5987 - 0.2266 = 0.3721 \quad \square$$

We want to approximate binomial probabilities of events like:

$$\{\text{from } k \text{ to } \ell \text{ successes in } n \text{ trials}\}$$

using the normal dist'n.

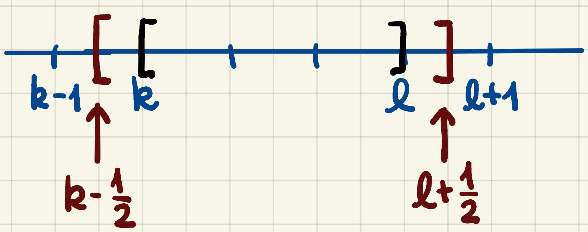
We set: $\mu = n \cdot p$

$\sigma = \sqrt{n \cdot p \cdot (1-p)}$

Our idea for the approximation is

$$\boxed{\Phi\left(\frac{\ell - \mu}{\sigma}\right) - \Phi\left(\frac{k-\mu}{\sigma}\right)}$$

$\Rightarrow$ We resort to the so-called continuity correction.



$\mathbb{P}[\{\text{from } k \text{ to } \ell \text{ successes in } n \text{ trials}\}] =$

Integers

$$= \mathbb{P}[\{\text{from } k-\tfrac{1}{2} \text{ to } \ell+\tfrac{1}{2} \text{ successes in } n \text{ trials}\}] \approx$$

$$\approx \Phi\left(\frac{\ell + \frac{1}{2} - \mu}{\sigma}\right) - \Phi\left(\frac{k - \frac{1}{2} - \mu}{\sigma}\right)$$

Then, for $k = \ell$:

$$\Phi\left(\frac{k-\mu}{\sigma} + \frac{1}{2\sigma}\right) - \Phi\left(\frac{k-\mu}{\sigma} - \frac{1}{2\sigma}\right)$$

Problem. Defective Objects.

Each object produced by a factory is defective w/ probability 0.1.
Assume that objects are independent.

Q: What is the distribution of the number of defective objects among 200 chosen @ random from the assembly line?

$\longrightarrow$: Binomial$(n=200, p=0.1)$.

Q: What is the expression for the exact probability that @ most 20 of the 200 objects are defective?

$\longrightarrow$:
$$\sum_{k=0}^{20}\binom{200}{k}(0.1)^k(0.9)^{200-k}$$

Q: What is the approximate probability of the above event?

$\longrightarrow$: $\mathbb{P}$[the # of defective objects is ≤ 20] = $\boxed{\text{by convention}}$

$=\mathbb{P}$[between $-\infty$ and 20 "successes"]

We now use the normal approximation:

$$\mu = 200(0.1) = 20$$
$$\sigma = \sqrt{20(0.9)} = \sqrt{18} = 3\sqrt{2}$$

$$\Phi\left(\frac{\cancel{20}+\frac{1}{2}-\cancel{20}}{3\sqrt{2}}\right) - \underbrace{\Phi(-\infty)}_{=0} = \Phi\left(\frac{1}{6\sqrt{2}}\right)$$

Method I: pnorm(1/6/sqrt(2)) = 0.5469072

1/(6*sqrt(2))

Method II: $\Phi(0.117851) = \Phi(0.12) = 0.5478$

0.12

Q: When is the normal approximation applicable?

$\boxed{\text{When } n\cdot p \geq 10 \text{ and } n(1-p) \geq 10}$