Stat 401 12/4/17

Sec 5.3 Continued

Recall: CLT

Let $X_1, X_2, \ldots, X_n$ denote the observations of a random sample from a distribution that has mean μ and finite variance σ^2.

Then $\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1)$.

i.e. $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$

i.e. $\bar{X}_n \xrightarrow{D} N\left(\mu, \frac{\sigma^2}{n}\right)$

if start w/ dist. that's not normal, always have this convergence (approximation).

if start w/ dist that IS normal, this is exact.

ex) Let $\bar{X}$ denote the mean of a random sample of size 75 from the distribution that has pdf

$$f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{o.w.} \end{cases}$$

$$\mathbb{P}\{0.45 < \bar{X} < 0.55\}.$$

$\boxed{\text{SOL}}$ $\mathbb{E}[X] = \int_0^1 x \cdot 1 \, dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} = \mu$

$$\mathbb{E}[X^2] = \int_0^1 x^2 \cdot 1 \, dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

$$\text{Var}[X] = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12} = \sigma^2$$

$$\mathbb{P}\left\{\frac{0.45 - 0.5}{\sqrt{\frac{1/12}{75}}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{0.55 - 0.5}{\sqrt{\frac{1/12}{75}}}\right\}$$

$$= \mathbb{P}\{-1.5 < Z < 1.5\}$$

$$= \Phi(1.5) - \Phi(-1.5)$$

$$= 0.9332 - 0.0668$$

$$= 0.8664$$

ex) Suppose $X_1, \ldots, X_n$ is a random sample from a distribution with mean μ and unknown variance σ^2.

Then we can say

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \xrightarrow{D} N(0,1)$$

S = sample std dev.

Proof)

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\sigma}{S} \cdot \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

Now $S^2 \xrightarrow{P} \sigma^2$ (sec 5.1)

$$\Rightarrow S \xrightarrow{P} \sigma$$

$$\Rightarrow \frac{S}{\sigma} \xrightarrow{P} 1$$

By CLT $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$.

Recall Slutsky's Thm.

$$\text{If } X_n \xrightarrow{D} X, \; B_n \xrightarrow{P} b$$

$$\text{then } B_n X_n \xrightarrow{D} bX$$

$$\therefore \frac{\sigma}{S} \cdot \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1).$$

Normal Approximation to Binomial Dist

Suppose $X_1, \ldots, X_n$ is a random sample where $X_i \sim Ber(p)$.

$[\mu = p, \; \sigma^2 = p(1-p)]$

If $Y_n = X_1 + \cdots + X_n$ we know

$$Y_n \sim Bin(n, p)$$

$$\text{By CLT } \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1).$$

$$\Rightarrow \frac{\sum X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1)$$

$$\frac{\sum X_i - n\mu}{\sigma\sqrt{n}} = \frac{\sum X_i - np}{\sqrt{p(1-p)}\sqrt{n}}$$

$$= \frac{Y_n - np}{\sqrt{np(1-p)}}$$

$Y_n = X_1 + \cdots + X_n$

$$= \frac{n\bar{X}_n - np}{\sqrt{np(1-p)}}$$

$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$

$$= \frac{n(\bar{X} - p)}{\sqrt{np(1-p)}}$$

$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$

$$= \frac{\sqrt{n}(\bar{X} - p)}{\sqrt{p(1-p)}} \xrightarrow{D} N(0,1)$$

Limiting dist of Binomial RV is Normal w/mean 0 variance 1

$$Y_n \xrightarrow{D} N(np, np(1-p))$$

$$\bar{X}_n \xrightarrow{D} N\left(np, \frac{p(1-p)}{n}\right)$$

This approx. works well when

$np \geq 5$ and $n(1-p) \geq 5$.

Note: The area of the rectangle whose base is $(k-0.5,\ k+0.5)$ and the area under the normal curve between $k-0.5$ and $k+0.5$ are approx. equal.

called a continuity correction.

ex) $Y \sim \text{Bin}(100, \frac{1}{2})$

Find $\mathbb{P}[\underbrace{Y = 48, 49, 50, 51, 52}_{\text{event is equivalent to the event}}]$.

SOL

event is equivalent to the event

$$\{47.5 < Y < 52.5\}.$$

$\mu = np = 100(\tfrac{1}{2}) = 50$

$\sigma^2 = np(1-p) = 100(\tfrac{1}{2})(\tfrac{1}{2}) = 25$

$\sigma = 5$

$\mathbb{P}[47.5 < Y < 52.5]$

$= \mathbb{P}\left[\frac{47.5-50}{5} < Z < \frac{52.5-50}{5}\right]$

$= \mathbb{P}[-0.5 < Z < 0.5]$

$= \phi(0.5) - \phi(-0.5)$

$= 0.6915 - 0.3085$

$= 0.3830$

ex) Prob. a patient recovers from a disease is 0.4. Suppose 100 people are known to have this disease.

a) What is prob. that less than 30 survive?

$Y \sim Bin(100, 0.4)$

$P[Y < 30] = P[Y < 29.5]$.

$\{0, 1, \ldots, 29\}$

$\approx P\left[Z < \frac{29.5 - 40}{\sqrt{24}}\right]$

$\approx P[Z < -2.14]$

≈ 0.0162

using Binomial 0.0148

b) $P[Y \leq 40] = P[Y < 40.5]$

$\{0, \ldots, 40\}$ $\approx P\left[Z < \frac{40.5 - 40}{\sqrt{24}}\right]$

$\approx P[Z < 0.10]$

≈ 0.5398

binomial 0.5433

$P[1 < Y < 30] = P[1.5 < Y < 29.5]$

$P[Y > 50] = P[Y > 50.5]$