Statistics, lecture 17:- يوتيوب 22 + 23

*The normal approximation to the binomial distribution.

في الكتاب ، اعلى قيمة ل n في ال binomial هي 25 بالكثير
اذا اعطاك Bin وال n>25 ، لازم تعرف انك تقرب من طريقه ال normal .
لازم الشرطين يتحققوا عشان اقرب
$np \geq 5$ and $nq \geq 5$

*If n is large and P is small or moderate then let $X \sim Bin(n,p)$ Then $X \mathrel{\dot\sim} n(np, npq)$

Continuity correction:-

*If $X = 1$, $X \in (0.5, 1.5)$

*If $X = 5$, $X \in (4.5, 5.5)$

e.g) If $X \sim Bin(100, 0.2)$, find approximate values for the following probabilities:-

i) $P(X < 26)$

ii) $P(X \leq 26)$

iii) $P(18 < X \leq 26)$

iv) $P(18 \leq X < 26)$

v) $P(18 \leq X \leq 26)$

sol) $n = 100$ $P = 0.2$ $q = 0.8$

$\mu = np = 20$ $\sigma^2 = npq = 16$

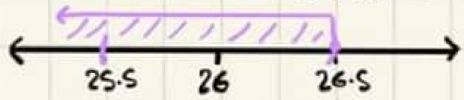
$X \mathrel{\dot\sim} n(20, 16)$

i) $P(X < 26)$

$P(X \leq 25.5) = P(Z \leq 1.38)$

↳ وضع مساواة او لا ؟ مارح تفرق X لما صارت normal $= 0.9162$

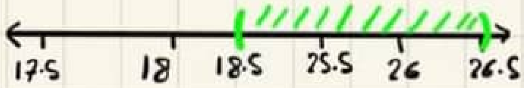
ii) $P(X \leq 26) =$ المساواة دون تأثير



$P(X \leq 26.5) = P(Z \leq 1.63)$

مساواة او بدون ما تفرق ع 25 ↲ $= 0.9484$

iii) $P(18 < X \leq 26)$



$P(18.5 \leq X \leq 26.5) =$

$P(-0.38 \leq Z \leq 1.63) =$

$P(Z \leq 1.63) - P(Z \leq -0.38)$

$0.9484 - 0.3520$

$= 0.5964$

ملاحظة : لما نحسب $\frac{18.5-20}{4}$ بكون الجواب -0.375 لكن قربناها ل -0.38

iv) $P(18 \leq X < 26)$

$P(17.5 \leq X \leq 25.5) = P(-0.63 \leq Z \leq 1.38)$

$P(Z \leq 1.38) - P(Z \leq -0.63)$

$0.9162 - 0.2643$

$= 0.6519$

v) $P(18 \leq X \leq 26)$

$P(17.5 \leq X \leq 26.5) = P(-0.63 \leq Z \leq 1.63)$

$P(Z \leq 1.63) - P(Z \leq -0.63)$

$= 0.9484 - 0.2643$

$= 0.6841$

e.g) suppose that 10% of heavy smokers will suffer from lung cancer after the age of 40, in a sample of 100 heavy smokers, what is the Prob that:

a) at least 12 will have lung cancer.

b) no more than 14 will have lung cancer.

c) Exactly 12 will have lung cancer.

Sol)

$n = 100 \quad P = 0.10 \quad q = 0.90$

$\mu = 10 \quad \sigma^2 = 9 \quad \sigma = 3$

$X \sim n(10, 9)$

a) $P(X \geq 12) =$

$P(X \geq 11.5) = 1 - P(Z \leq 0.5)$

$= 1 - 0.6915$

$= 0.3085$

b) $P(X \leq 14) =$

$P(X \leq 14.5) = P(Z \leq 1.5)$

$= 0.9332$

c) $P(X = 12) =$

$P(11.5 \leq X \leq 12.5) = P(0.5 \leq Z \leq 0.83)$

$P(Z \leq 0.83) - P(Z \leq 0.5)$

$0.7967 - 0.6915$

$= 0.1052$

****

* The central limit theorem (C.L.T).

If a random sample $X_1, X_2, \ldots X_n$ is drawn from a population with mean μ and variance σ^2, then for large n, the distribution of sample mean $\bar{X}$ is approximately normal with mean μ and variance $\frac{\sigma^2}{n}$:-

$\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$ for $n \geq 30$

*مق بستخدم هاي الطريقة لمّا يسألني عن $\bar{X}$ او مجموع قيم X .

e.g) Suppose that a random sample of $n=100$, is drawn for population with mean 70 and standard deviation 20, what is the Prob. that the sample mean will be:

a) more than 70.

b) less " 73.

sol) $n = 100 \quad \mu = 70 \quad \sigma^2 = 400 \quad \sigma = 20$

$X_1, X_2, X_3 \ldots X_{100} \; n \sim (70, 400)$

↳ Random sample = Independent.

$\bar{X} \sim n(70, 4)$

a) $P(\bar{X} > 70) = 1 - P(Z \leq 0)$

$= 1 - 0.5000$

$= 0.5$

b) $P(\bar{X} < 73) = P(Z \leq 1.5)$

$= 0.9332$

e.g) Suppose that the mean weight and standard deviation of orange boxes are 10 and 2 kgs respectively. If 100 boxes are to be loaded in a car with threshold 1000 kgs what is the Prob. that the car will break down ؟

Sol) $X_1, X_2, X_3, \ldots \sim n(10, 4)$

$\bar{X} \sim n(10, 0.04)$ بيسأل عن احتمالية

$P(\sum_{i=1}^{100} X_i > 1000) = P(\bar{X} > 10)$ انه السيارة تبطل

$= 1 - P(Z \leq 0)$ تشتغل، عشان

0.5 تبطل تشتغل لازم

يكون مجموع وزن الصناديق > 1000 بالتالي المتوسط > 10

ك الحمولة القصوى "Threshold"

e.g) solve the previous example if the threshold of the car is 1050 kg.

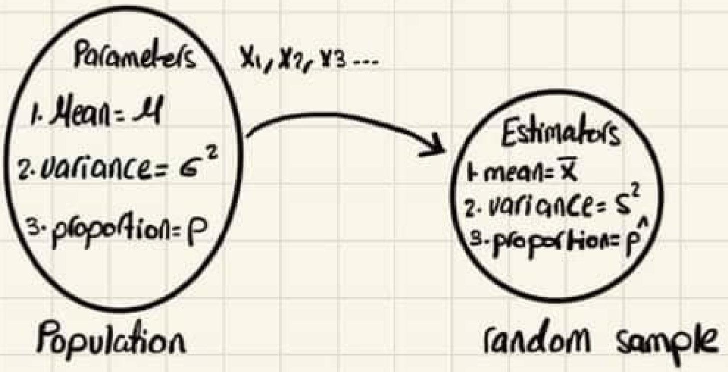
$P(\sum_{n=1}^{n=100} X_i > 1050) = P(\bar{X} > 10.5)$

$= 1 - P(Z < 2.5)$

$= 1 - 0.9938$

$= 0.0062$

* Sampling distribution *



Sampling distribution: نهتم بتوزيع ال estimators

* The distribution of the sample mean *

If $X_1, \ldots, X_n$ is a r.s "random sample" from a population with mean μ and variance σ^2, then

a) the mean of $\bar{X}$ is $\mu \Rightarrow E(\bar{X}) = \mu$

b) " variance of $\bar{X}$ is $\frac{\sigma^2}{n} \Rightarrow var(\bar{X}) = \frac{var(X)}{n}$

c) If $X_1, \ldots, X_n \sim n(\mu, \sigma^2)$ then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$

e.g) suppose that the weights of a certain population are normally distributed with $\mu = 70$ kgs and standard deviation 10 kgs. If a sample of size $n = 25$ persons is to be drawn, what is the Prob. that:

a) the average weight will be less than 73 kgs.

b) " total weight exceeds 1800 kg.

Sol) $\mu = 70 \quad \sigma = 10 \quad \sigma^2 = 100$

$n = 25 \qquad X$: weight

$X_1, X_2, \ldots, X_{25} \sim n(70, 10^2)$

$\bar{X} \sim n(70, 2^2)$

a) $P(\bar{X} < 73) = P(Z < 1.5)$

$= 0.9332$

b) $P(\sum_{n=1}^{n=25} X_i > 1800) = P(\bar{X} > 72)$

$= 1 - P(Z \leq 1)$

$= 1 - 0.8413$

$= 0.1587$

Statistics: lecture 18 يوتيوب 25 + لقادمبا ش 12

* "The distribution of the sample proportion"

$P^\wedge = \frac{X}{n}$

$$P^\wedge \sim n\left(P, \frac{Pq}{n}\right)$$

or

$$Z = \frac{P^\wedge - P}{\sqrt{\frac{Pq}{n}}} \sim n(0,1)$$

وهون n دايماً large
بحفن التكر حكى أولاً

e.g) suppose that 10% of a certain product are defective. If 400 items are drawn from the production, what is the Prob. that the sample proportion will be:

a) more than 12%.

b) between 9% and 11%.

sol) $P = 0.10 \quad q = 0.90 \quad n = 400$

$P^\wedge \sim n(0.10, (\frac{3}{200})^2)$

a) $P(P^\wedge > 0.12) = 1 - P(Z \leq 1.33)$
$= 0.0918$

b) $P(0.09 < P^\wedge < 0.11) = P(-0.67 < Z > 0.67)$
$= 0.7486 - 0.2514$
$= 0.4972$

e.g) Suppose that 90% of the uni students Pass calculus 101, In a sample of 200 students taking calculus 101, what is the Prob that the proportion of those who will pass is less than 85%

sol) $P = 0.90 \quad q = 0.10 \quad n = 200$

$P \sim n(0.90, 0.00045)$

$P(P^\wedge < 0.85) = P(Z < -2.36)$
$= 0.0091$

Solve Q41+42, book, p296