Task 3

(a)

A government is considering the need for additional airport capacity due to projections of increased demand for air travel in the years ahead. The government is deciding how the expected demand can be met in the long term. Short-listed options for increasing airport capacity include expansion of one of two existing airports, A or B. The decision to expand either site involves numerous trade-offs.

You have been asked to devise an appropriate sampling scheme of airport A and airport B users (passengers as well as non-passengers, such as staff and local residents to the airport) to research their views of expanding one or other site. Explain how each of the following sampling methods could be applied to the overall sampling strategy for this study. Make sure you describe the merits and limitations of each as well as how each would be applied in practice.

Variant A

i. Convenience sampling.

ii. Quota sampling.

iii. Stratified sampling.

iv. Cluster sampling.

Variant B

i. Judgemental sampling.

ii. Snowball sampling.

iii. Systematic sampling.

iv. Cluster sampling.

– Questionaires
↳ non-prob. methods

– Data analysis
↳ prob. methods

A) i. Quest. in hall of airport

+ convenient
– not representative

ii.

Technique	Strengths	Weaknesses
Non-probability sampling		
Convenience sampling	Least expensive, least time-consuming, most convenient	Selection bias, sample not representative, not recommended for descriptive or causal research
Judgemental sampling	Low cost, convenient, not time-consuming ideal for exploratory research designs	Does not allow generalisation, subjective
Quota sampling	Sample can be controlled for certain characteristics	Selection bias, no assurance of representativeness
Snowball sampling	Can estimate rare characteristics	Time-consuming
Probability sampling		
Simple random sampling (SRS)	Easily understood, results projectable	Difficult to construct sampling frame, expensive, lower precision, no assurance of representativeness
Systematic sampling	Can increase representativeness, easier to implement than SRS, sampling frame not always necessary	Can decrease representativeness depending upon 'order' in the sampling frame
Stratified sampling	Includes all important subpopulations, precision	Difficult to select relevant stratification variables, not feasible to stratify on many variables, expensive
Cluster sampling	Easy to implement, cost effective	Imprecise, difficult to compute and interpret results

1) division by quotas
2) by convenience

1) division by stratas (cat. variable)
2) SRS within each strata

1) division by clusters
(mini - version of population)
2) SRS of cluster

E.g. : 1) Expert opinion on quota formation
2) Convenience sampling
+ control for main routes
− selection

iii. 1) Sort by frequency of the flights
↳ change to intervals
2) Proportional SRS

$$Route_i \times Freq_i \Rightarrow N_1 \times N_2$$
group

+ accurate
− costly & time - inefficient

iv. 1) Clusters = planes
2) Randomly select n planes
and question all people there

+ more convenient than
strat. sampling

- clusters can be not represantative

(b)

Suppose we are interested in estimating the mean of a population with a finite variance using a simple random sample of size n. $\rightarrow f(x_1, \ldots, x_n)$

i. State a suitable estimator of the population mean as well as its sampling distribution. Mention any assumptions which you make.

ii. Explain how the sampling distribution derived in i. should be interpreted.

iii. Explain how to determine the minimum sample size necessary to estimate a population mean to within e units assuming the population standard deviation is known. If the population standard deviation was unknown, how would you deal with this?

iv. Explain the purpose of the finite population correction factor (including a formula) and when it should be used.

i. Suitable :

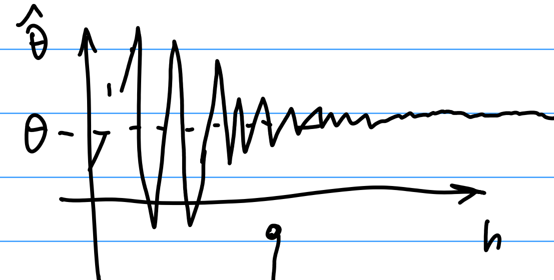
1) $E(\hat{\theta}) \xrightarrow{d} \theta \quad / \quad E(\hat{\theta}) = \theta$

2) $\hat{\theta} \xrightarrow{P} \theta$

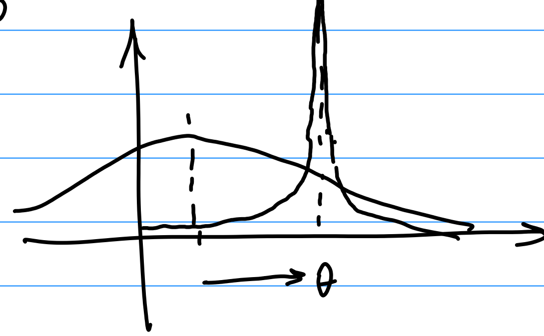
$\hookrightarrow E(\hat{\theta}) \xrightarrow{d} \theta$

$\oplus$

$Var(\hat{\theta}) \xrightarrow{d} 0$



3) min Var / MSE



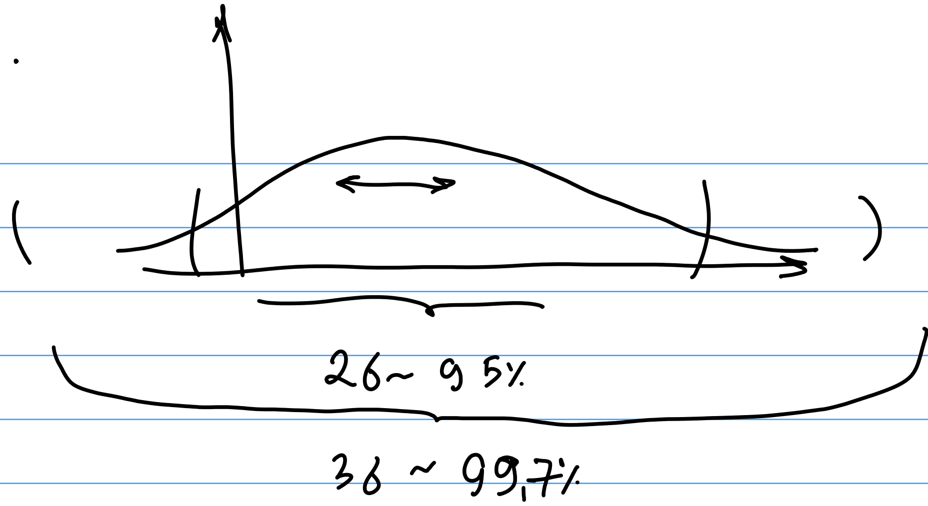
By CLT : $\bar{X} = \frac{1}{n} \sum X_i$ — estimator

$$\boxed{\bar{X} \overset{d}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)}$$

Assumption : for n large enough

$n \geq 20 \qquad X_i \sim N(\mu, \sigma^2)$

$n \geq 200 \qquad X_i$ from skewed, with heavy tails ii.



$$Var(\bar{x}) = f(n) \qquad n\uparrow \Rightarrow Var\downarrow$$

iii.

$$\{ \bar{x} \pm z_{1-\alpha/2} \cdot se(\bar{x}) \}$$

$$\{ \bar{x} \pm z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n}} \}$$

$$\underbrace{z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}}_{ME} \leq e$$

$$n \geq \frac{\sigma^2 \cdot z^2_{1-\alpha/2}}{e^2}$$

$$\left\{ \begin{array}{l} \alpha\downarrow \Rightarrow n\uparrow \\ e\downarrow \Rightarrow n\uparrow \\ \sigma^2\uparrow \Rightarrow n\uparrow \end{array} \right.$$

- MPC for industry
  for Russia

→ $s^2_x = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

→ from other analysis or literature

iv.

population $\to$ infinite

population $\to$ finite

rand. int

$\{X_1, \dots, X_n\}$

$\{X_1, \dots, X_1\}$

$\{X_2, X_4, \dots, X_2, X_4\}$

$N = 100$

$n = 20 \quad = \quad 20\%\, N$

e.g. $\overset{10}{K} \quad \overset{5}{3} \quad \overset{1}{C}$

$S = \{K\} \qquad p_K = 1/3$

$S = \{K; 3\} \qquad p_3 = 1/2$

Finite: $\quad cov(x_i, x_j) = -\dfrac{\sigma^2}{N-1}$

Inf.: $\quad cov(x_i, x_j) = 0 \qquad , N \to \infty$

$$\hat{\sigma}_{cor} = \hat{\sigma} \cdot \sqrt{\frac{N-n}{N-1}}$$

N - pop. size

n - sample size

$$F = \sqrt{\frac{N-n}{N-1}}$$

$N \to \infty, \quad n$ - fixed

$F = 1$

$$\sqrt{\frac{100-20}{100-1}} \approx 0,9 \implies \hat{\sigma}_c = \underbrace{0,9 \cdot \hat{\sigma}}_{corr} \quad \leftarrow conv.$$

CLT: $X_i \sim Be(p)$

$$\hat{p} \overset{d}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$$

$$X_i = \begin{cases} 0 & ; 1-p \\ 1 & ; p \end{cases} \qquad \bar{X} = \frac{\#1}{N} = \hat{p}$$

$$E(X_i) = p$$

$$\sqrt{\frac{p(1-p)}{n}} \cdot z_{1-\alpha/2} \leq e$$

$$Var(X_i) = p(1-p)$$

$\Rightarrow n$

$$se_c(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$