

NPU Nettverksmøte

Litt om forutsetninger for basestørrelser;
populasjon, utvalg og feilmarginer

Litt begrepsrydding

POPULASJON

Den totale gruppen av objekter/individer vi trekker vårt utvalg fra - for observasjon.

“den norske befolkning over 18 år”

UTVALG

Den delen av populasjonen vi trekker ut, og velger og observere.

Data fra utvalget er reelle og har ikke feilmarginer i seg selv



Litt begrepsrydding, forts.

REPRESENTATIVITET

Snakker vi om når vi ønsker å trekke konklusjoner om populasjonen, basert på observasjoner fra utvalget.

Et par regler:

1. Vi må kjenne trekksannsynligheten for alle objekter/individer for de variabler/kjennetegn vi definerer (oftest kjønn/alder/geografi).
I praksis - vi må ha data for alle objekter/individer!
2. Vi må trekke **tilfeldig** fra **populasjonen** vi skal undersøke.

FEILMARGINER/SIGNIFIKANS

Når vi bruker resultater fra et utvalg som resultater for en populasjon får vi det pr. definisjon aldri 100% presist - vi må leve med FEILMARGINER.

Signifikans dreier seg om å korrigere for feilmarginer når vi vurderer om resultater fra utvalget gjelder for populasjonen, så vi kan vurdere om en forskjell i utvalget kan antas å gjelde for hele populasjonen.



Feilmarginer

FEILMARGINER, TABELL

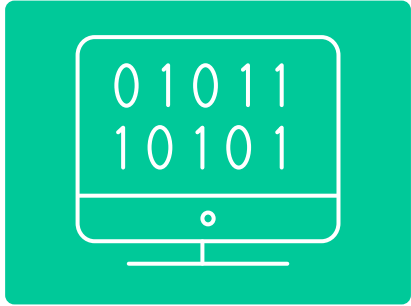
Tabellen for et rent lotterisk utvalg
(eks: NO-POPULASJON) angir feilmarginer
for ulike observasjoner ved ulike basestørrelser:

Antall observasjoner	Prosentresultat							
	1/99	5/95	10/90	15/85	20/80	30/70	40/60	50/50
50	2.8	6	8.3	9.9	11.1	12.7	13.6	13.9
100	2	4.3	5.9	7	7.8	9	9.6	9.8
150	1.6	3.5	4.8	5.7	6.4	7.3	7.8	8
170	1.5	3.3	4.5	5.4	6	6.9	7.4	7.5
200	1.4	3	4.2	4.9	5.5	6.4	6.8	6.9
300	1.1	2.5	3.4	4	4.5	5.2	5.5	5.7
335	1.1	2.3	3.2	3.8	4.3	4.9	5.2	5.4
400	1	2.1	2.9	3.5	3.9	4.5	4.8	4.9
500	0.9	1.9	2.6	3.1	3.5	4	4.3	4.4
600	0.8	1.7	2.4	2.9	3.2	3.7	3.9	4
778	0.7	1.5	2.1	2.5	2.8	3.2	3.4	3.5
800	0.7	1.5	2.1	2.5	2.8	3.2	3.4	3.5
813	0.7	1.5	2.1	2.5	2.7	3.2	3.4	3.4
900	0.7	1.4	2	2.3	2.6	3	3.2	3.3
1000	0.6	1.4	1.9	2.2	2.5	2.8	3	3.1
1600	0.5	1.1	1.5	1.7	2	2.2	2.4	2.5
3000	0.4	0.8	1.1	1.3	1.4	1.6	1.8	1.8
5000	0.3	0.6	0.8	1	1.1	1.3	1.4	1.4
5347	0.3	0.6	0.8	1	1.1	1.2	1.3	1.3
6500	0.2	0.5	0.7	0.9	1	1.1	1.2	1.2

Prosentresultat pluss/minus feilmarginen i tabellen angir et intervall som med 95% sannsynlighet inneholder det sanne resultatet



Basestørrelser, en oversikt



BRUKERGRUPPEN

Kjenner vi den demografiske fordelingen på våre brukere? Kjønn, alder og evt. geografi

Dette er et KRAV for å kunne vurdere robustheten i basestørrelser

KONSISTENS

For en brukergruppe med ukjent demografi trenger vi tid og konsistent overvåking for å vurdere hvordan gruppen er fordelt, og hvordan dette varierer (gjennom).

Dette gir grunnlaget vi trenger

FORMÅL

Det er stor forskjell på om vi trenger en "pekepinn", eller ønsker robuste svar på viktige spørsmål.

Sjekk de viktigste demografiske fordelingene!

BASESTØRRELSE

Vi vurderer i praksis sjelden baser under 30-40 respondenter.

Baser med mer enn 100 respondenter kan gi gode svar dersom vi kjenner populasjonen

Så hva er egentlig konklusjonen på vurdering av basestørrelser?

- Bruk tid på å sjekke hvordan brukergruppen ser ut
- Vi må vite mest mulig om våre brukere før vi kan vurdere om en base er stor nok
- Når vi har tilstrekkelig informasjon kan vi vurdere om en undergruppe ligner på vår totale brukergruppe (eller på en hvilken annen gruppe vi ønsker å sammenligne oss med)
- Vær oppmerksom på at ulike events kan ha svært forskjellige brukertyper
- Ikke vurder baser under 30 respondenter
- **Spør gjerne!**

Know your world

We deliver data to trust, for your most well-founded decisions.

