

Forprosjekt – Statistikk

Gruppediskusjon om et stokastisk forsøk

Verdimengde

Verdimengden til en stokastisk variabel er mengden av alle mulige verdier variabelen kan anta. Skostørrelse er en diskret variabel, og basert på våre data blir verdimengden:

- $X = \{42, 43, 47\}$.

Høyde er i teorien en kontinuerlig variabel, men siden vi kun har en endelig mengde observerte verdier, blir verdimengden:

- $Y = \{173, 183, 186, 189, 193\}$

Diskret/kontinuerlig

I teorien vil høyde være kontinuerlig og skostørrelse være diskret ettersom at skostørrelser er faste intervaller mens høyde kan i teorien være uendelig detaljert. I rammene av vårt forsøk vil derimot både og høyde og skostørrelse være diskret ettersom at vi avrunder høyden.

Forventningsverdi

Teoretisk sett vil forventningsverdien være gjennomsnittlig verdi for skostørrelse/høyde for norsk person i aldersgruppe student. I kontekst av vårt datasett vil forventningsverdien bare bli gjennomsnittets verdi fra datasettet vårt.

Standardavvik

Standardavvik gir oss ett tall på hvor stor spredning vi har mellom dataene. Hvis det er ett stort standardavvik vet vi at det er stor variasjon mellom dataene våre og gjennomsnittshøyden/skostørrelsen, og ett lite standardavvik vil si at alle høydene/skostørrelsene er veldig nære gjennomsnittet.

Kovarians

Ettersom at høyde korrelerer med skostørrelse, vil X og Y ha en kovarians. Da forventes det en positiv kovarians. Det vil si at når X øker, øker Y også, punktene danner en stigende sky, og regresjonslinja vil ha positiv helning.

Form på sannsynlighetsfordelingene for X og Y

X (skostørrelse) er en diskret heltallsvariabel. I en blandet NTNU-populasjon forventer vi en svakt bimodal fordeling (en "klokke" for kvinner og en for menn (kirke klokke eller bellcurve (på engelsk))). Innen hver delgruppe er den unimodal og nær normal form rundt en typisk størrelse.

Y (høyde) er også en diskret heltallsvariabel. For voksne er høyde tilnærmet normalfordelt innen hver delgruppe (én symmetrisk klokke). Sammenslått gruppe kan gi to overlappende klokker (bimodal).