

# Deskriptiv statistik

Population - ofta oändlig.

Om hela populationen känd, så behövs ej ML.

## Stickprov

ett stickprov med storlek  $n$  från en fördelning  $X$  är en samling av  $n$  oberoende slumpvariabler, alla med fördelningen  $X$ .

Om  $n \leq 5\%$  av populationen, så är det ok utan återläggning, men i allmänhet måste urvalet ske med återläggning.

# Statistika

En statistika är en slumpvariabel som är en funktion av ett stickprov.

Ex:

Stickprovsmedlet,  $\bar{X} \neq \mu$  skrivs  $\bar{X}$  oftast

$$\frac{\sum_{i=1}^n x_i}{n}$$

Detta är ett lägesmått (location statistic)

## Fler lägesmått

Median, medianläge (median location)

Medianläget i ett sorterat stickprov finns på platsen

$\frac{n+1}{2}$ , där  $n$  är storleken på stickprovet

Om vi har "mellan" tal är medianen medlet av värdet ovanför och under medianläget.

Typvärde (mode)

Det mest förekommande värdet.

Statistikor för spridning (variability)  
min, max (range)

Varians

$$\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1} = S^2$$

Standard avvikelsern ( $\sigma$ ):

$$\sqrt{S^2} = S$$

uppskattning av  $\sigma$ , inte  
pop. värde

Om vi har hela populationen

då: 
$$\sigma^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n}$$