

ST0103 Brukerkurs i statistikk

Høst 2016

Forelesning 4, 2.9.2016

Regel 3.13 Multiplikasjonsregelen

La A og B være to hendelser definert i et utfallsrom S . Hva er sannsynligheten for at både A og B inntreffer?

sannsynligheten for A og B

= sannsynligheten for A \times sannsynligheten for B gitt A ,

dvs.

$$P(A \text{ og } B) = P(A) \cdot P(B|A)$$

...

eller også (siden A og B kan byttes om:)

$$P(A \text{ og } B) = P(B) \cdot P(A|B)$$

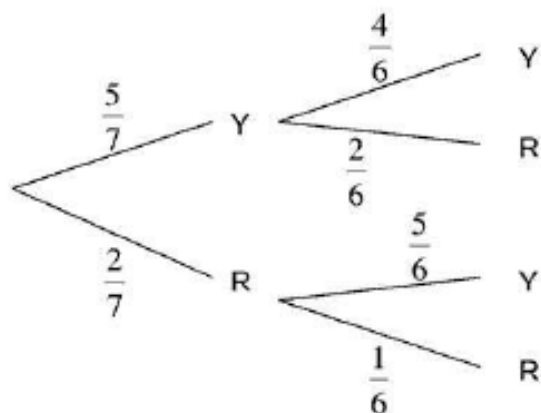
Trekking uten tilbakelegging

En bolle inneholder 7 kuler, 5 gule (Y) og to røde (R). To kuler trekkes uten tilbakelegging, dvs. at det først trekkes en og at det så trekkes en til uten å legge den første tilbake. La

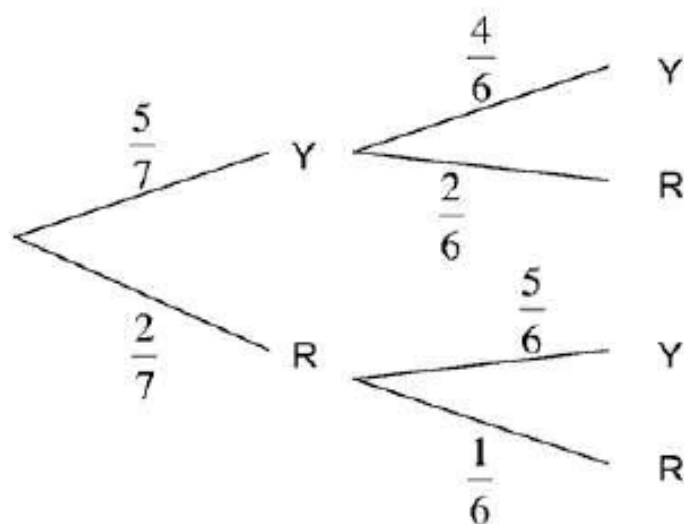
A = den første kulan er gul (Y)

B = den andre kulan er gul (Y)

$$P(\text{begge kulene er gule}) = P(A \text{ og } B) = P(A) \cdot P(B|A) = \frac{5}{7} \cdot \frac{4}{6} = \frac{20}{42}$$



Forts.



$$P(\text{en av hver farge}) = P(YR) + P(RY) = \frac{5}{7} \cdot \frac{2}{6} + \frac{2}{7} \cdot \frac{5}{6} = \frac{10}{42} + \frac{10}{42} = \frac{20}{42}$$

$$P(\text{to røde}) = P(RR) = \frac{2}{7} \cdot \frac{1}{6} = \frac{2}{42}$$

Diagnostiske tester

- S = syk person, \bar{S} = frisk person.
- T = positiv test, \bar{T} = negativ test.
- For legemidler “vet” man:
 - $P(T|S)$: Sannsynligheten for at testen slår ut positivt, gitt at personen er syk (*sensitiviteten* til testen). Ønskes høyest mulig.
 - $P(\bar{T}|\bar{S})$: Sannsynligheten for at testen slår ut negativt, gitt at personen er frisk (*spesifisiteten*). Ønskes høyest mulig.
- Interessant for pasienten:
 - $P(S|T)$: Sannsynligheten for at du er syk, gitt at du har fått en positiv test (*positiv prediktiv verdi*).
 - $P(\bar{S}|\bar{T})$: Sannsynligheten for at du er frisk, gitt at du har fått en negativ test (*negativ prediktiv verdi*).

Hvorfor utføres ikke HIV-test som masseundersøkelse?



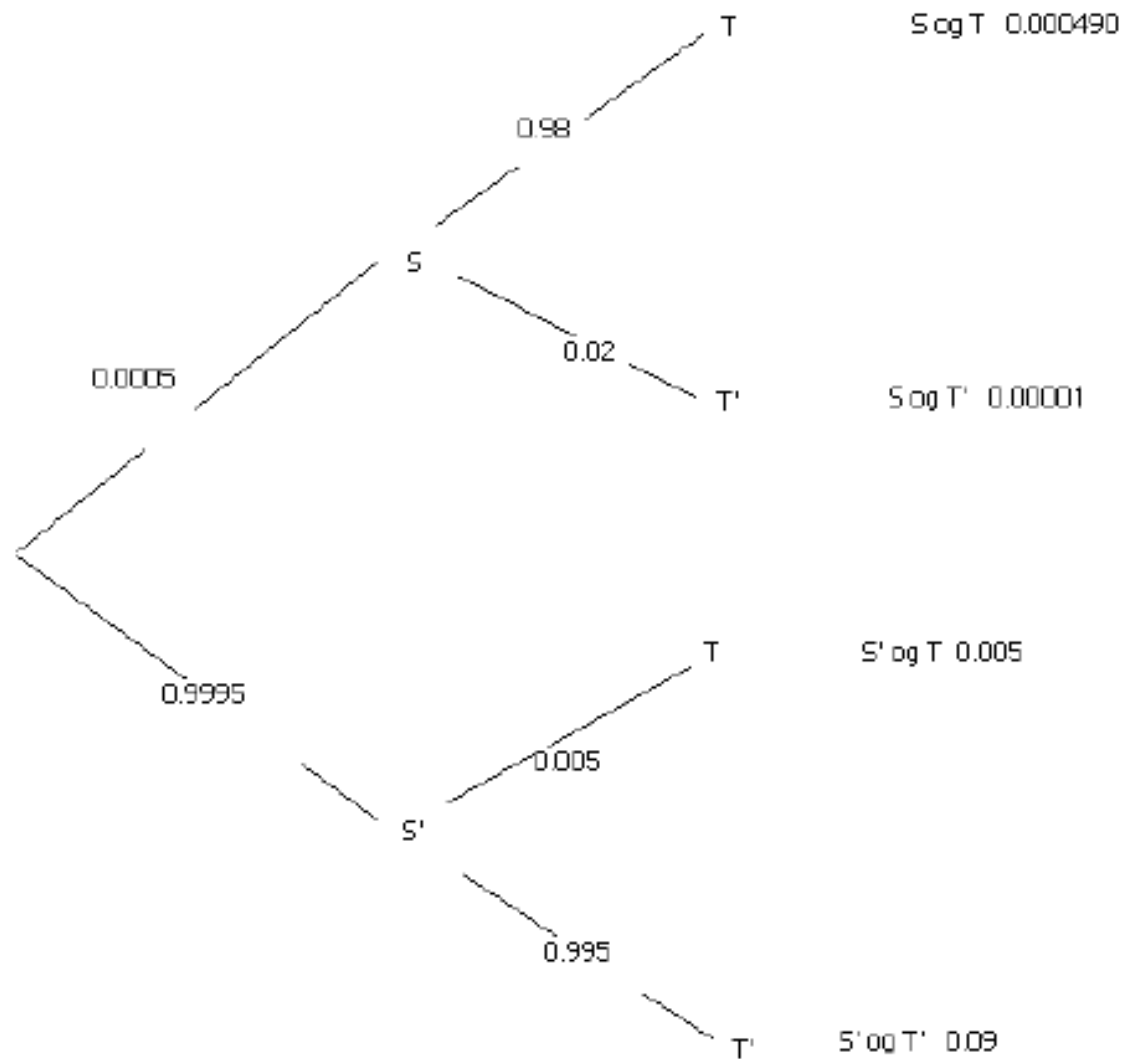
$$P(S|T) = \frac{P(S \text{ og } T)}{P(T)}$$

— Hva er sannsynligheten for at en person med positiv HIV-test virkelig er HIV-smittet, $P(S|T)$?
Anta

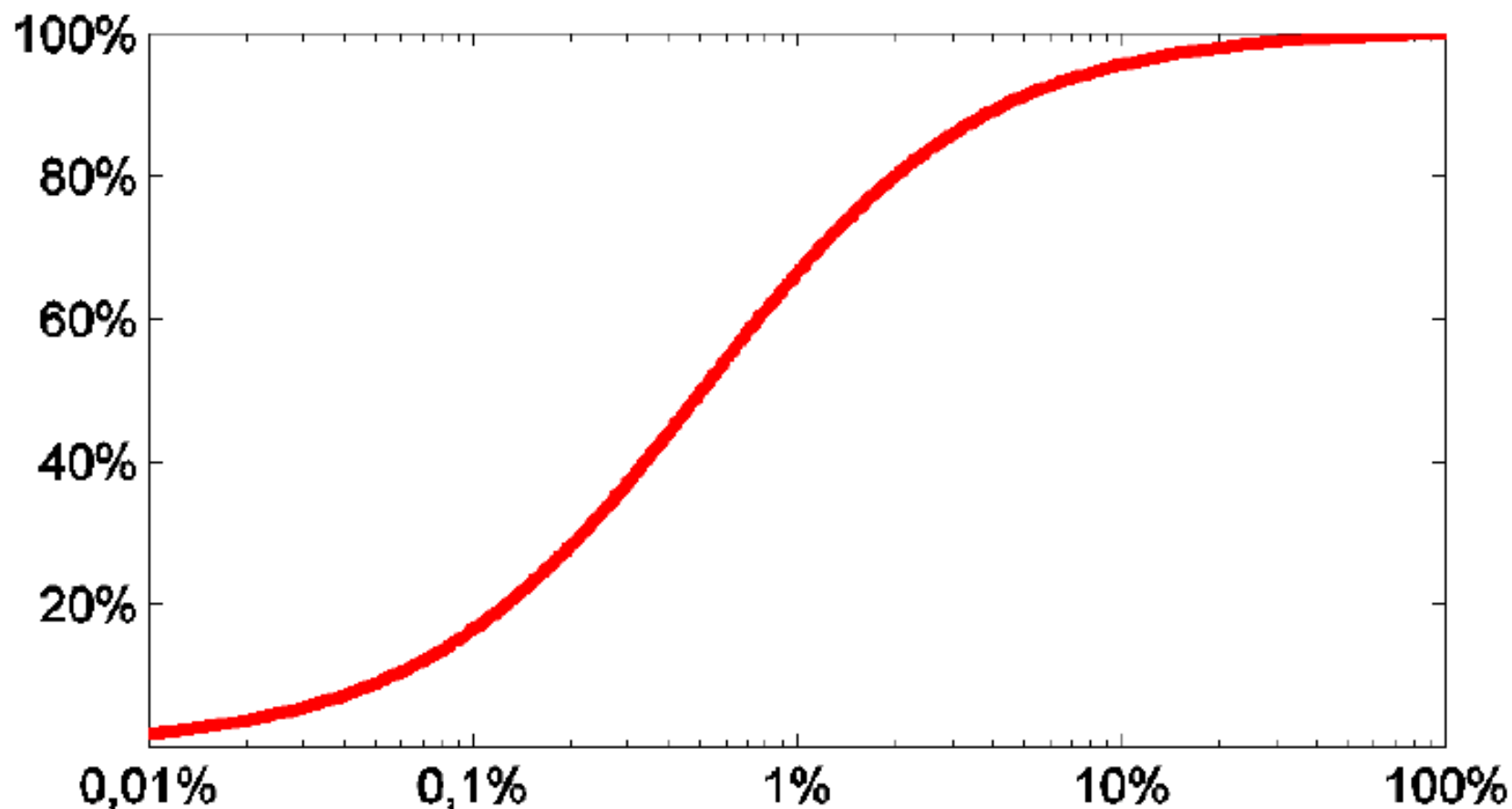
- Sensitivitet av testen:
 $P(T|S) = 0.98$
- Spesifisitet av testen:
 $P(\bar{T}|\bar{S}) = 0.995$, dvs.
 $P(T|\bar{S}) = 0.005$.

— Svaret er avhengig av forekomsten av HIV i populasjonen, $P(S)$ (prevalensen).

— Anta at forekomsten av HIV i en populasjon er $P(S) = 0.0005$.



$$P(S|T) = \frac{P(S \text{ og } T)}{P(T)} = \frac{0.000490}{0.000490 + 0.005} = 0.09$$



x-akse: andel smittede i befolkningen, $P(S)$

y-akse: andel som er smittet blant de med positiv test, $P(S|T)$

HIV-test



- Norge som helhet:
 $P(S) = 0.001$ (anslag fra lærebok i medisinsk statistikk) gir $P(S|T) = 0.16$.
- Sprøytemisbrukere: $P(S) = 0.1$ gir $P(S|T) = 0.956$.
- Storby i sentral-Afrika:
 $P(S) = 0.25$ gir
 $P(S|T) = 0.985$.

— Dette gir et problem ved masseundersøkelser. De fleste av personene med positiv prøve kan faktisk være friske!