Alla medicinare bör känna till resonemanget kring Bayes' sats

När är ett positivt test sant positivt?

ATTILA FRIGYESI, fil dr, specialistläkare i kardiologi, ST-läkare anestesi och intensivvård, Universitetssjukhuset i Lund attila@frigyesi.se

För att inte skrämma iväg alltför många läsare med sannolikhetslära vill jag börja med ett exempel som skulle kunna förekomma i den kliniska vardagen. Antag att vi har ett test för t ex HIV med 99 procents sensitivitet (om man undersöker HIV-bärare med detta test kommer man att i medel få positivt utfall i 99 fall av 100) och 99 procents specificitet (om man undersöker icke-HIV-bärare med detta test kommer man att i medel få negativt utfall i 99 fall av 100). I klinisk praxis är detta ett mycket bra test. Ska vi då testa alla i Grönköping med detta test? Låt oss anta att prevalensen är 0,5 procent. Vi har alltså följande situation:

- sensitivitet 99 procent
- specificitet 99 procent
- prevalens 0,5 procent.

Vad är då sannolikheten för HIV-bärarskap om vi har positivt test?

A) 90 procent
B) 50 procent
C) 33 procent.

Intuitivt skulle man kunna tycka att sannolikheten borde bli ganska hög, eftersom vårt utmärkta test utföll positivt. Skulle det vara 33 procent är det mer sannolikt att personen inte är HIV-bärare än att personen vore det! I anglosaxisk litteratur kallas ibland prevalensen för pre-test probability, medan sannolikheten för HIV om vi har ett positivt test kallas för post-test probability. Vi återkommer strax till detta illustrativa exempel.

Så räknar man fram Bayes' sats

Låt oss som brukligt beteckna sannolikheten för händelsen HIV-bärarskap med P(HIV). Antingen är man HIV-bärare eller så är man det inte, alltså P(HIV) + P(ej HIV) = 1, varur följer P(HIV) = 1 – P(ej HIV).

Vi behöver införa begreppet betingad sannolikhet. Om man undersöker HIV-bärare med ett test kan man få positivt utfall (+) eller negativt utfall (–). Givet att man är HIV-bärare så betecknas sannolikheten för positivt utfall P(+|HIV). Detta är inget annat än vårt tests sensitivitet. Hur definieras då denna sannolikhet? Antag att det i Grönköping finns 20 000 personer och att alla dessa testats med vårt test och att man vet vilka som är HIV-bärare:

	HIV	ej HIV
+	99	199
–	1	19 701

Vårt tests sensitivitet är:

$$\frac{\text{antal(HIV och +)}}{\text{antal(HIV)}} = \frac{99}{99+1} = 0{,}99$$

Om vi nu dividerar både täljare och nämnare med antalet invånare i Grönköping fås:

$$\frac{\frac{\text{antal(HIV och +)}}{20\,000}}{\frac{\text{antal(HIV)}}{20\,000}} \approx \frac{\text{P(HIV och +)}}{\text{P(HIV)}}$$

Alltså är en bra definition av den betingade sannolikheten (i vårt fall testets sensitivitet)

$$\text{P(+|HIV)} = \frac{\text{P(HIV och +)}}{\text{P(HIV)}}$$

Med andra beteckningar fås

$$\text{P(A|B)} = \frac{\text{P(A och B)}}{\text{P(B)}}$$

som kan skrivas om till

$$\text{P(A och B)} = \text{P(A|B)} \times \text{P(B)}$$

Om man noterar att P(A och B) är precis samma sak som P(B och A) så får man

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

Dividera båda leden med P(B). Då fås Bayes' sats:

$$\text{P(A|B)} = \frac{\text{P(B|A)} \times \text{P(A)}}{\text{P(B)}}$$

SAMMANFATTAT

Med Bayes' sats kan man, med känd sensitivitet, specificitet och sjukdomsprevalens, beräkna sannolikheten för att ett positivt testresultat är sant positivt.

Eftersom prevalensen ofta är låg, kommer även ett bra test att medföra att ett positivt resultat mest sannolikt är falskt positivt. Man bör alltså inte förlita sig på ett enkelt test.

Det är bra att kunna Bayes' sats, men viktigast är att ha en känsla för ovanstående resonemang.

Om det är så att antingen händelsen A eller ej A och inget annat inträffar (exempelvis HIV-bärare eller icke-HIV-bärare), kan man skriva Bayes' sats i formen

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B|A) \times P(A) + P(B|ej\ A) \times P(ej\ A)}$$

där nämnaren enligt definitionen för betingad sannolikhet blir P(B och A) + P(B och ej A) = P(B).

För den intresserade läsaren kan Gunnar Bloms »Sannolikhetsteori med tillämpningar« [1] och Johan Brings och Adam Taubes »Introduktion till medicinsk statistik« [2] rekommenderas.

Förvånande sannolikhet

Åter till vårt exempel. Vi undrade över sannolikheten/risken, om vi fått ett positivt testutfall hos en slumpmässigt vald Grönköpingsbo, att denna person är HIV-bärare, symboliskt P(HIV|+). Vi känner till sensitivitet och specificitet, dvs P(+|HIV) och P(–|ej HIV). Om vi nu använder Bayes' sats får vi

$$P(HIV|+) = \frac{P(+|HIV) \times P(HIV)}{P(+|HIV) \times P(HIV) + P(+|ej\ HIV) \times P(ej\ HIV)}$$

Nu är det bara att fylla i:

$$P(HIV|+) = \frac{0{,}99 \times 0{,}005}{0{,}99 \times 0{,}005 + 0{,}01 \times 0{,}995} \approx 0{,}33$$

Något förvånande är alltså sannolikheten för att vår person som testades positiv för HIV verkligen är HIV-bärare 0,33, vilket innebär att det är mer sannolikt att vår person är icke-HIV-bärare!

Orsaken till detta är att även om vårt test är specifikt, kommer 1 procent av dem som inte är HIV-bärare att testas positivt, vilket ger litet eftertanke. Eftersom icke-HIV-bärarna är så många fler än HIV-bärarna (99,5 procent mot 0,5 procent), kommer majoriteten av dem som testas positivt faktiskt att vara falskt positiva.

Vi ska inte lita på ett enda test

Detta exempel illustrerar det viktiga faktum att vi inte ska lita på ett enda test, i synnerhet när det gäller sjukdomsentiteter med låg prevalens.

Behöver då en medicinare kunna Bayes' sats? Man behöver kanske inte kunna Bayes' sats, men resonemanget kring detta viktiga exempel bör alla medicinare känna till, eftersom vi ofta ställs inför diagnostiska test av olika slag. Man bör vara försiktig med att testa stora grupper individer med låg sjukdomsprevalens, s k screening. Man bör aldrig förlita sig på ett enstaka test!

Som kuriosa kan nämnas att Bayes' sats har fått sitt namn från den brittiske matematikern och presbyterianska kyrkoherden Thomas Bayes, som levde 1709–1761. Han formulerade ett specialfall av Bayes' sats, vilket publicerades postumt [3].

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Kommentera denna artikel på lakartidningen.se

REFERENSER

1. Blom G. Sannolikhetsteori med tillämpningar. Lund: Studentlitteratur; 1998.
2. Bring J, Taube A. Introduktion till medicinsk statistik. Lund: Studentlitteratur; 2006.
3. Price R. Essay towards solving a problem in the doctrine of chances. London: Philosophical Transactions of the Royal Society of London; 1764.