Statistika v kazenskem pravu

Dolga predstavitev diplomske naloge

Neža Kržan

Mentor: izr. prof. Jaka Smrekar
Fakulteta za matematiko in fiziko

Ljubljana, 28.november 2022

Kratek pregled

Kratek pregled

Kratek pregled

Koncept verjetnosti

Definicija

Naj bo H nek dogodek in $\bar{H}$ negacija oziroma komplement dogodka H. Dogodka H in $\bar{H}$ sta znana kot komplementarna dogodka.

Opravlja se primerjava verjetnosti dokazov na podlagi dveh konkurenčnih predlogov:
$H_p \ldots$ trditev, ki jo predlaga tožilstvo;
$H_d \ldots$ trditev, ki jo predlaga obramba.

Koncept verjetnosti je

- ključen pri ocenjevanju dokazov;
- omogoča objektivno oceno vpliva dokaza na verjetnost določene domneve o interesni osebi ali obdolžencu.

Presoja dokazov

- Za presojo se uporablja različne metode in tehnike, ki temeljijo na statistični verjetnosti,
- metode omogočajo oceno, kako verjetno je, da so dokazi resnični in zanesljivi.

KONTEKST DOKAZA

- Upoštevamo druge dokaze in okoliščine primera.
- Izvedemo bolj objektivno presojo dokazov.
- Vpliv na določeno domnevo oziroma hipotezo o interesni osebi ali obdolžencu.
- Upoštevamo verjetnost napake - verjetnost, da so dokazi napačni ali zavajajoči.

Koncept verjetnosti

Verjetnost proti nedolžnosti ali verjetnost za krivdo

$$\frac{P(H_p)}{P(H_d)}$$

Verjetnost v prid krivdi ob upoštevanju informacij E

$$\frac{P(H_p|E)}{P(H_d|E)}$$

Če imamo na voljo dokaz E, nas zanima pogojna verjetnost

$$P(kriv|E),$$

pri čemer nam je lahko v pomoč Bayesovo pravilo.

Opredelitev

Bayesovo sklepanje razlaga verjetnost kot merilo verjetnosti ali zaupanja, ki ga lahko ima posameznik glede nastanka določenega dogodka.

- O dogodku lahko že imamo predhodno prepričanje oziroma apriorno prepričanje.
- Ta se lahko spremeni, ko se pojavijo novi dokazi.

Bayesova statistika nam daje matematične modele za vključevanje naših apriornih prepričanj in dokazov za ustvarjanje novih prepričanj.

Bayesova analiza

standardna metoda za posodabljanje verjetnosti po opazovanju več dokazov, zato je zelo primerna za sintezo dokazov.

- Začnemo z nekim predhodnim prepričanjem o hipotezi,
- posodabljamo to prepričanje, ko se dokazi ponovno pojavijo.

Pri uporabi Bayesovega sklepanja moramo utemeljiti predhodne predpostavke, kadar je to mogoče.
V nasprotnem primeru morajo uporabiti razpon vrednosti predpostavk in analizo občutljivosti, da preverijo zanesljivost rezultata. glede na te vrednosti.

Bayesovo pravilo

Bayesovo sklepanje temelji na Bayesovem pravilu, ki izraža verjetnost nekega dogodka z verjetnostjo dveh dogodkov in obrnejnje pogojne verjetnosti.

Definicija

Pogojna verjetnost dogodka H, glede na dogodek E, je

$$P(H|E) = \frac{P(H \cap E)}{P(E)}, \tag{1}$$

ob predpostavki, da je $P(E) > 0$.

Bayesovo pravilo

Potem pogojno verjetnost uporabimo še v števcu formule (1) in dobimo Bayesovo pravilo

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)},$$

verjetnost dogodka E lahko še razpišemo

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E|H)P(H) + P(E|\neg H)P(\neg H)}.$$

Obstaja še ena formulacija Bayesovega pravila, ki olajša izračune in je pogosto uporabljena pri Bayesovi analizi DNK dokazov

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{P(E|H)}{P(E|\neg H)} \times \frac{P(H)}{P(\neg H)}.$$

Bayesovo posodabljenje

Logična trditev, kako se sčasoma posodabljajo apriorne oziroma predhodne verjetnosti dokazov glede na novo zbrane dokaze oziroma prepričanja.

Trditev (Bayesovo posodabljanje)

Če se dogodek E zgodi ob času $t_1 > t_0$, potem je $P_1(H) = P_0(H|E)$.

- Predhodna oz. apriorna verjetnost = $P_0(H)$ (ob času t_0 dogodku H dodelimo verjetnost);
- Ko se zgodi dogodek E ob času t_1, ki vpliva na naša prepričanja o dogodku H, Bayesovo posodabljanje pravi, da je potrebno apriorno verjetnost dogodka H v času t_1 enačiti z pogojno verjetnostjo dogodka H glede na dogodek E v času t_0.

Predhodna verjetnost in določitev posteriorne verjetnosti

Recimo, da je sodni izvedenec je naprošen, da opravi analizo profila DNK krvi, najdene na kraju kaznivega dejanja, in rezultat primerja s profilom DNK obdolženca.

Odločitev porotnikov bo delno odvisna od njihove ocene dveh interesnih hipotez
$H_1 \ldots$ vir krvi je obtoženec,
$H_2 \ldots$ vir krvi je druga oseba.

Porotniki bodo morda želeli, da jim dokončno povemo, katera hipoteza je resnična, ali da jim navedemo verjetnosti vira. Za oceno verjetnosti vira mora forenzični znanstvenik upoštevati tudi druge dokaze v kazenskem primeru.

Predhodna verjetnost in določitev posteriorne verjetnosti

Obtoženec in kri s kraja zločina imata skupen niz t.i. genetskih označevalcev - najdemo pri eni osebi na 1 milijon prebivalcev.

$\rightarrow$ **pogojna verjetnost ugotovitve rezultatov pri dveh hipotezah o medsebojni povezanosti**

- skupni genetski označevalci skoraj zagotovo najdeni v primeru H_1 (vir je bil obtoženec);
- 1 možnost na milijon, da bodo najdeni v primeru H_2 (vir je bil nekdo drug);

$\rightarrow$ **predložimo razmerje verjetnosti**

Predhodna verjetnost in določitev posteriorne verjetnosti

Edini skladen način, kako na podlagi forenzičnih dokazov sklepati o verjetnosti virov, je uporaba Bayesovega pravila, ki zahteva, da začnemo s pripisom predhodnih verjetnosti za hipoteze, ki nas zanimajo.

Bayesov pristop

Deluje le, če bomo lahko začeli s predhodno oziroma apriorno verjetnostjo.

Predhodna verjetnost in določitev posteriorne verjetnosti

Ali naj analitiki poskušajo določiti predhodne verjetnosti in če ja, kako naj jih določijo?

- **Nevtralno stanje** - analitiki predpostavijo enake predhodne verjetnosti za vse hipoteze v primeru;

PRAKTIČNA IN OBJEKTIVNA ANALIZA

- uporabimo Bayesovo pravilo za izračun posteriorne verjetnosti;

Predpostavljanje enake verjetnosti za vse možnosti je problematično, kajti v realnosti se različne hipoteze razlikujejo po svoji verjetnosti.

Poenostavljena Bayesova analiza

Naj bo:
S ... trditev, da je obtoženec vir sledi DNK s kraja zločina;
M ... trditev, da se obtoženčev DNK ujema z DNK-jem s kraja zločina;
f ... funkcija pogostosti ujemanja DNK z DNK-jem s kraja zločina.
Želimo vedeti, kakšna je verjetnost S glede na M, tj. $P(S|M)$.

Bayesovo pravilo lahko uporabimo na naslednji način:

$$\frac{P(S|M)}{P(\neg S|M)} = \frac{P(M|S)}{P(M|\neg S)} \times \frac{P(S)}{P(\neg S)}.$$

Poenostavljena Bayesova analiza

$$\frac{P(S|M)}{P(\neg S|M)} = \frac{P(M|S)}{P(M|\neg S)} \times \frac{P(S)}{P(\neg S)}.$$

Verjetnosti $P(S)$ in $P(\neg S)$ je težko oceniti, ker ne vemo kakšna je množica osumljencev;

$P(M|S) = 1$ - **lažno negativni rezultat**;

$P(M|\neg S) = f$ - **lažno pozitivni rezultat**;

Ob upoštevanju lažno negativnega in pozitivnega rezultata sledi:

$$\frac{P(S|M)}{P(\neg S|M)} = \frac{1}{f} \times \frac{P(S)}{P(\neg S)}.$$

Izpopolnjena Bayesova analiza I

Za upoštevanje možnosti laboratorijskih napak, bomo namesto M uvedli spremenljivko M_p.

M_p . . . poročano ujemanje laboratorijske analize;
M_t . . . trditev, da obstaja dejansko ujemanje v DNK-ju;
$\neg M_t$ . . . trditev, da obstaja neujemanje v DNK-ju.

Sledi:

$$P(M_p|\neg S) = P(M_p|M_t)P(M_t|\neg S) + P(M_p|\neg M_t)P(\neg M_t|\neg S).$$

Izpopolnjena Bayesova analiza I

$$P(M_p|\neg S) = P(M_p|M_t)P(M_t|\neg S) + P(M_p|\neg M_t)P(\neg M_t|\neg S).$$

Sedaj je:

- $P(M_t|\neg S) = f$ in zato
- $P(\neg M_t|\neg S) = 1 - f$;

$P(M_p|\neg M_t) = FP$ ...verjetnost lažno pozitivnih rezultatov laboratorija;

$P(M_p|M_t) = FN$ ...verjetnost lažno negativnih rezultatov laboratorija;

Sledi:

$$P(M|\neg S) = [(1 - FN) \times f] + [FP \times (1 - f)].$$

Izpopolnjena Bayesova analiza II

Predpostavimo, da je

$$P(M_p|S) = 1,$$

pri čemer ni upoštevana možnost lažnega negativnega rezultata.

$$P(M_p|S) = P(M_p|M_t)P(M_t|S) + P(M_p|\neg M_t)P(\neg M_t|S)$$

$P(M_p|S) = 1 \to P(\neg M_p|S) = 0$ in $P(M_p|S) = 1 - FN$

$$\frac{P(M_p|S)}{P(M_p|\neg S)} = \frac{1 - FN}{[(1 - FN) \times f] + [FN \times (1 - f)]}$$

Izpopolnjena Bayesova analiza

- Relativno majhne stopnje napak lahko bistveno zmanjšajo dokazno vrednost DNK dokazov, saj močno zmanjšajo razmerje verjetnosti.
- Vpliv stopnje laboratorijskih napak kaže, da ne glede na to, kako nizka se izkaže pogostost profila, bo ta relativno nepomembna, če pogostosti ne spremlja ocena stopnje laboratorijskih napak.

Spremenljivost pogostosti profila lahko v Bayesovem okviru upoštevamo na dva načina:

1. s spremembo predhodne verjetnosti;
2. s spremembo pogostosti profila.