

Pri opredelitvi področja uporabe statističnih podatkov s področja kazenskega pravosodja je pomembno vedeti, kdo so pretekli in sedanji uporabniki ter za kakšen namen se bodo podatki uporabljali. Vprašanja se lahko nanašajo tudi na to, kdo predloži podatke, kdo jih prejme, kateri podatki se predložijo ter v kakšni obliki in v kakšnih časovnih presledkih. Pri obravnavi samega sistema kazenskopravne statistike je treba razmisliti tudi o tem, katera ključna politična vprašanja je treba vključiti v oblikovanje vsakega programa za njegovo izboljšanje.

Za dober statistični model v pravosodju, bi morali za izhodišče upoštevati naslednje vidike:

- pojavnost kaznivih dejanj (resnost, trendi, struktura, drugo),
- profil storilcev kaznivih dejanj in njihove značilnosti,
- delovna obremenitev sistema (kazniva dejanja, aretacije, izreki kazni, storilci kaznivih dejanj pod nadzorom),
- storilci kaznivih dejanj in primeri, ki se gibljejo po sistemu,
- značilnosti žrtev,
- porabljena sredstva (človeška in finančna),
- korelacije kriminala, kot so gospodarski, demografski in drugi podatki,
- družbeni in gospodarski stroški kaznivih dejanj,
- odnos državljanov do kriminala in kazenskega pravosodja ter njihova zaskrbljenost v zvezi z njima

Države zbirajo ogromne količine podatkov, ki se nanašajo na kazniva dejanja, kriminalne profile in z njimi povezane socialno-ekonomske, politične in geografske profile različnih skupnosti in države kot celote, da bi podprle svoj kazenskopravni sistem. V mnogih primerih vlade za pridobivanje podatkov izvajajo obsežne raziskave s kvantitativnimi ali kvalitativnimi tehnikami. Med raziskovalnim procesom je potrebna statistična analiza, da se vsi ti podatki spremenijo v uporabne informacije za učinkovito odločanje na področju kazenskega pravosodja.

## 1. VREDNOTENJE DOKAZOV

**1.1. Komplementarni dogodki (angl. Complementary Events).** Naj bo  $R$  nek dogodek in  $\bar{R}$  negacija oziroma komplement dogodka  $R$ . Dogodka  $R$  in  $\bar{R}$  sta znana kot komplementarna dogodka. Pogosto bom opravila primerjavo verjetnosti dokazov na podlagi dveh konkurenčnih predlogov, in sicer predloga tožilca in predloga obrambe.

$H_p \dots$  trditev, ki jo predlaga tožilstvo;

$H_d \dots$  trditev, ki jo predlaga obramba;

Oznaka  $p$  in  $d$  označujeta obtožnico (angl. proposition) oziroma obrambo (angl. defence), črka  $H$  pa označuje hipotezo.

Hipoteze se lahko dopolnjujejo na enak način, kot dogodki. Ena in samo ena je

lahko resnična; med seboj se izključujejo. Ni nujno, da so izbrane tako, da zajemajo vse možne razlage dokazov. Dve hipotezi lahko označujeta komplementarne dogodke (npr. resnično kriv in resnično nedolžen), vendar pa se lahko zgodi, da se označena dogodka ne dopolnjujeta.

**1.2. Koncept verjetnosti.** Koncept verjetnosti je pomemben pri ocenjevanju dokazov, saj se le ti ocenjujejo glede na njihov vpliv na verjetnost določene domneve o interesni osebi (v nadaljevanju PoI) (preden pride do sojenja) ali obdolžencu (medtem ko sojenje poteka). Zanima nas vpliv dokazov na verjetnost krivde ( $H_p$ ) in nedolžnosti ( $H_d$ ) osumljenca. Gre za dopolnjujoča se dogodka in razmerje verjetnosti teh dveh dogodkov,

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

je verjetnost proti nedolžnosti ali verjetnost za krivdo. Ob upoštevanju dodatnih informacij  $I$ , je razmerje

$$\frac{P(H_p|I)}{P(H_d|I)},$$

verjetnost v prid krivdi ob upoštevanju informacij  $I$ .

Občasno se zgodi, da predloga tožilstva in obrambe nista komplementarna in v takih primerih ni mogoče določiti  $P(H_p)$  ali  $P(H_d)$ , ampak samo vpliv statistike, znane kot razmerje verjetnosti (angl. Likelihood Ratio; oznaka LR).

**1.3. Bayesova teorija.** Bayesova analiza je standardna metoda za posodabljanje verjetnosti po opazovanju več dokazov, zato je zelo primerna za sintezo dokazov. Vsakdo, ki mora presoditi o hipotezi, kot je "krivda" (vključno s preiskovalci pred sojenjem, sodniki, porotami), neformalno začne z nekim predhodnim prepričanjem o hipotezi in ga posodablja, ko se dokazi ponovno pojavijo. Včasih lahko obstajajo celo objektivni podatki, na katerih temelji predhodna verjetnost. Bayesovo sklepanje je veljavna metoda za izračun posteriorne verjetnosti. Pri uporabi Bayesovega sklepanja morajo statistiki utemeljiti predhodne predpostavke, kadar je to mogoče, na primer z uporabo zunanjih podatkov; v nasprotnem primeru morajo uporabiti razpon vrednosti predpostavk in analizo občutljivosti, da preverijo zanesljivost rezultata glede na te vrednosti.

Bayesov izrek za dogodka  $S$  in  $R$ :

$$P(S|R) = \frac{P(R|S)P(S)}{P(R)},$$

kjer  $P(R) \neq 0$ .

## 2. VERJETNOSTNA OBLIKA BAYESOVEGA IZREKA

**2.1. Razmerje verjetnosti (angl. Likelihood Ratio).** V zgornji formuli nadomestimo  $S$  z  $\bar{S}$  in enakovredna različica Bayesovega izreka je

$$P(\bar{S}|R) = \frac{P(R|\bar{S})P(\bar{S})}{P(R)},$$

kjer  $P(R) \neq 0$ .

Če prvo enačbo delimo z drugo dobimo verjetnostno obliko Bayesovega izreka

$$\frac{P(S|R)}{P(\bar{S}|R)} = \frac{P(R|S)}{P(R|\bar{S})} \times \frac{P(S)}{P(\bar{S})}.$$

Leva stran je verjetnost dogodka  $S$  ob pogoju, da se je zgodil dogodek  $R$ . Pogojna verjetnost na desni strani dogodka,  $S$  in  $\bar{S}$ , sta v števcu in imenovalcu različni, medtem ko je dogodek  $R$ , katerega verjetnost nas zanima, enak. Na koncu pa imamo verjetnost v korist dogodka  $S$  brez kakršnih koli informacij o  $R$ .

Razmerje

$$\frac{P(R|S)}{P(R|\bar{S})}$$

je pomembno pri vrednotenju dokazov in se imenuje razmerje verjetnosti (angl. Likelihood Ratio) ali Bayesov faktor (angl. Bayes' factor, oznaka BF).

Oglejmo si dogodka  $R$  in  $S$ , ter njuni dopolnitvi. Razmerje verjetnosti je tu razmerje verjetnosti  $R$ , ko je  $S$  resničen in verjetnosti  $R$ , ko je  $S$  neresničen. Da bi upoštevali učinel  $R$  na verjetnost  $S$ , tj. da bi  $\frac{P(S)}{P(\bar{S})}$  spremenili v  $\frac{P(S|R)}{P(\bar{S}|R)}$ , prvo pomnožimo z razmerjem verjetnosti. Verjetnost  $\frac{P(S)}{P(\bar{S})}$  je znana kot predhodna verjetnost v korist  $S$ , verjetnost  $\frac{P(S|R)}{P(\bar{S}|R)}$  pa je znana kot posteriorna verjetnost v korist  $S$ . Razlika med  $P(R|S)$  in  $P(S|R)$  je bistvena.

Pri preučevanju vpliva  $R$  na  $S$  je treba upoštevati tako verjetnost  $R$ , ko je  $S$  resničen in ko je  $S$  neresničen. Pogosta napaka (zmota prenesene pogojne verjetnosti) je, da dogodek  $R$ , ki je malo verjeten, če je  $\bar{S}$  resničen, pomeni dokaz v prid  $S$ . Da bi bilo tako, je treba dodatno zagotoviti, da  $R$  ni tako malo verjeten, če je  $S$  resničen. Razmerje verjetnosti je potem večje od 1 in pozitivna verjetnost je večja od predhodne verjetnosti.

**2.2. Bayesov faktor (angl. Bayes' Factor) in razmerje verjetnosti (angl. Likelihood Ratio).** V forenziki se, kljub pogostejši uporabi Bayesovega faktorja (BF), pojma pogosto obravnavata kot sinonima. Bayesov faktor je glavni element Bayesove metodologije za primerjavo konkurenčnih predlogov. Opredeljen je kot sprememba, ki jo povzročijo novi dokazi (podatki) v verjetnosti pri prehodu od predhodne k posteriorni porazdelitvi v korist enega predloga k drugemu. Da se pokazati, da je razmerje verjetnosti poseben primer Bayesovega faktorja, kadar so konkurenčne hipoteze parametrizirane z enim samim parametrom (tj. preprosta hipoteza). Vendar pa lahko pride do primerov, ko se primerjajo sestavljene hipoteze. V takem primeru je Bayesov faktor razmerje dveh mejnih verjetnosti pri konkurenčnih hipotezah in se zdi, da ni več odvisen samo od podatkov.

**2.3. Logaritem razmerja verjetnosti.** Verjetnost in razmerje verjetnosti imata vrednosti med 0 in  $\infty$ , logaritmi teh statistik zajemajo vrednosti na  $(-\infty, \infty)$ . Verjetnostna oblika Bayesovega izreka vključuje multiplikativno razmerje, če pa vzamemo logaritme, postane razmerje aditivno:

$$\log \left( \frac{P(S|R)}{P(\bar{S}|R)} \right) = \log \left( \frac{P(R|S)}{P(R|\bar{S})} \right) + \log \left( \frac{P(S)}{P(\bar{S})} \right).$$

Zamisel o vrednotenju dokazov z dodajanjem logaritma k predhodni verjetnosti je v skladu z intuitivno zamisljo o tehtanju dokazov na tehtnici pravičnosti, zato je logaritem razmerja verjetnosti dobil tudi ime teža dokazov (angl. The Weight of evidence).

Razmerje verjetnosti z vrednostjo, večjo od 1, ki vodi k povečanju verjetnosti v korist  $S$ , ima pozitivno utež. Razmerje verjetnosti z vrednostjo, manjšo od 1, ki zmanjšuje verjetnost za  $S$ , ima negativno utež. Pri razmerju verjetnosti, katerega vrednost je enaka 1, je verjetnost v korist  $S$ , tehtnica pa ostane nespremenjena. Dokazi so logično pomembni le, če se verjetnost, da se ti dokazi ujemajo ob upoštevanju resničnosti nekaterih trditev, razlikuje od verjetnosti, da se isti dokazi ujemajo ob upoštevanju neresničnosti sporne trditve, tj. razmerje logaritemske verjetnosti ni enako nič. Logaritem razmerja verjetnosti oz. utež zagotavlja enakovredno merilo pomembnosti. Ta metoda je ugodna, ker izenačuje pomembnost dokazov, ki jih ponujata tako tožilstvo kot obtoženec. Ima lastnosti simetrije in aditivnosti. Matematično simetrijo med težo dokazov za tožilstvo in težo dokazov za obtoženca je mogoče ohraniti tako, da se pri obravnavi predloga obtoženca teža dokazov obrne.

### 3. VREDNOST DOKAZOV

**3.1. Vrednotenje forenzičnih dokazov.** Obravnavajmo obliko Bayesovega izreka o verjetnosti v forenzičnem kontekstu ocenjevanja vrednosti nekaterih dokazov. Naj bo:

$H_p \dots$  interesna oseba (PoI) oz. obtoženec je resnično kriv - nadomestimo  $S$ ;

$H_d \dots$  interesna oseba (PoI) je resnično nedolžen - nadomestimo  $\bar{S}$ ;

$Ev \dots$  obravnavani dokaz - nadomestimo dogodek  $R$ ;

To sedaj lahko zapišemo kot  $(E, M) = (E_c, E_s, M_c, M_s)$ , vrsta dokaza in njegova opazanja, kot je opisano na začetku. Oblika Bayesovega izreka nato omogoča, da se prehodne verjetnosti (tj. pred predstavitvijo  $Ev$ ) v korist krivde posodobijo v posteriorne verjetnosti ob upoštevanju  $Ev$ , na naslednji način:

$$\frac{P(H_p|Ev)}{P(H_d|Ev)} = \frac{P(Ev|H_p)}{P(Ev|H_d)} \times \frac{P(H_p)}{P(H_d)}.$$

Ob upoštevanju informacij o ozadju  $I$ , dobimo zapis

$$\frac{P(H_p|Ev, I)}{P(H_d|Ev, I)} = \frac{P(Ev|H_p, I)}{P(Ev|H_d, I)} \times \frac{P(H_p|I)}{P(H_d|I)}.$$

Pri vrednotenju dokazov  $Ev$  sta potrebni dve verjetnosti - verjetnost dokazov, če je PoI kriv in glede na informacije o ozadju, ter verjetnost dokazov, če je PoI nedolžen in glede na informacije o ozadju. Informacije o ozadju so včasih znane kot okvir okoliščin ali pogojne informacije.

Da lahko ocenimo oziroma določimo vrednost dokaza potrebujemo razmerje verjetnosti (angl. Likelihood ratio).

**Definicija 3.1.** Naj bosta  $H_p$  in  $H_d$  dve konkurenčni hipotezi ter  $I$  informacije o ozadju. Vrednost  $V$  dokaza  $Ev$  je podana z

$$V = \frac{P(Ev|H_p, I)}{P(Ev|H_d, I)},$$

razmerje verjetnosti, ki pretvori predhodne verjetnosti  $\frac{P(H_p|I)}{P(H_d|I)}$  v posteriorne verjetnosti  $\frac{P(H_p|Ev,I)}{P(H_d|Ev,I)}$ .

Kvalitativna lestvica za poročanje o vrednosti  $V$  podpore dokazov za  $H_p$  proti  $H_d$  (Vir: ENFSI, 2015)

TABELA 1. Kvalitativna lestvica za poročanje o vrednosti  $V$  podpore dokazov za  $H_p$  proti  $H_d$  (Vir: ENFSI, 2015).

1	$< V \leq$	2	brez podpore
2	$< V \leq$	10	šibka podpora prvi hipotezi
10	$< V \leq$	100	zmerna podpora prvi hipotezi
100	$< V \leq$	1000	srednje močna podpora prvi hipotezi
1000	$< V \leq$	10000	močna podpora prvi hipotezi
10000	$< V \leq$	1000000	zelo močna podpora prvi hipotezi
1000000	$< V$		izjemno močna podpora prvi hipotezi

3.1.1. *Utemeljitev uporabe razmerja verjetnosti.* Verjetnost hipoteze  $H$  na podlagi nekega dokaza  $E$  je verjetnost, da najdemo  $E$ , če je  $H$  resnična. Za alternativno hipotezo je LR razmerje obeh verjetnosti. LR nam pove, katera hipoteza je bolj podprta z dokazi. Kadar sta hipotezi medsebojno izključujoči in izčrpni, nam LR pove več. V tem primeru, če je verjetnost  $H$  večja od verjetnosti alternative, lahko sklepamo tudi, da se verjetnost  $H$  zaradi najdbe  $E$  poveča, medtem ko se verjetnost alternative zmanjša. Če je le mogoče, je treba upoštevati verjetnosti za vse razumne alternativne hipoteze (tako da je nabor hipotez izčrpen). Če se obravnavajo samo nekatere hipoteze, je treba pojasniti, da so predstavljene samo LR za pare teh hipotez. V primerih, ko je treba združiti več hipotez in/ali več dokazov, se lahko LR bolje uporablja v povezavi z drugimi metodami.

Kadar je treba količinsko ovrednotiti skupni učinek več dokazov, ki vključujejo različne povezane hipoteze (kot so hipoteze o ravni vira, ravni dejavnosti in ravni kazni-vega dejanja), poenostavljene rešitve, ki neupravičeno predpostavljajo neodvisnost, niso ustrezne. Grafični prikazi dokazov so lahko v veliko pomoč pri modeliranju odvisnosti. Obstaja interaktivna programska oprema za izvajanje izračunov na grafičnih modelih (Bayesovih mrežah), ki uporabnikom omogoča, da raziščejo vplive različnih predpostavk. Čeprav je takšne metode težko uvesti neposredno na sodišču, so koristne za sintezo dokazov v kateri koli fazi preiskave pred sojenjem.

Ocena vrednosti razmerja verjetnosti je lahko podvržena številnim virom negotovosti, vključno s kakovostjo podatkov, pridobljenih z analizami, ki jih opravijo forenzični znanstveniki, izbiro kontrolnega vzorca in najdenih predmetov, ki jih lahko vzamejo različni preiskovalci ali analizirajo različni analitiki ali laboratoriji. Ocena znanstvenih dokazov na sodišču pogosto zahteva kombinacijo podatkov o pojavu ciljnih značilnosti skupaj z osebnim poznavanjem okoliščin iz določenega primera. Jasno je, da ima vsaka ocena verjetnosti, ki se nanaša na določen primer, tudi če jo obravnavamo v obliki frekvence, sestavino, ki temelji na osebnem znanju. Drugi viri negotovosti vključujejo pridobivanje predhodnih verjetnosti, pogojenih z razpoložljivim znanjem, ali celo izvajanje numeričnih postopkov za razreševanje računskih težav. Poročilo o vrednosti razmerja verjetnosti vključuje merilo njegove natančnosti, na primer z navedbo številčnega razpona vrednosti za verjetnost dokazov na

podlagi konkurenčnih predlogov in s tem številčnega razpona vrednosti za razmerje verjetnosti. Vendar sta vrednost dokaza in moč posameznikovega prepričanja o vrednosti različna pojma in se ne smeta združevati v intervalu ali povzročiti spremembe vrednosti dokaza, kot se to na primer zgodi z navedbo spodnje meje neke poljubno izbrane ravni. V praksi je za kriminalistično preiskavo na voljo en niz podatkov o ozadju, ki so značilni za člane določene relevantne populacije, en niz kontrolnih podatkov in en niz izterjanih podatkov. Zato je za vrednotenje dokazov z določenim statističnim modelom na voljo ena sama vrednost  $V$  za povezano razmerje verjetnosti. Ponovno je treba upati, da so vsi različni kontrolni vzorci in pridobljeni podatki dovolj reprezentativni za populacije, iz katerih so bili izbrani, tako da se bodo razmerja verjetnosti po vrednosti le malo razlikovala.

**3.2. Vloga informacij o ozadju.** Iz zgornjih enačb je razvidno, da je vrednost dokaza odvisna od osnovnih informacij  $I$  ter hipotez  $H_p$  in  $H_d$ . Osnovne informacije so osebne za vsakega posameznika. Zato je vrednost dokaza za vsakega posameznika drugačna, posteriorne verjetnosti so potem za vsakega posameznika drugačne in zato vrednotenje dokazov nima smisla. Dva udeleženca v sodnem postopku označimo kot A in B, informacije o ozadju, ki so na voljo vsakemu od njiju, pa kot  $I_a$  oziroma  $I_b$ , pri čemer so skupne informacije o ozadju, ki so jima na voljo,  $I = I_a \cup I_b$ . Bayesov izrek potem lahko zapišemo kot

$$\frac{P(H_p|Ev, I_a \cup I_b)}{P(H_d|Ev, I_a \cup I_b)} = \frac{P(Ev|H_p, I_a \cup I_b)}{P(Ev|H_d, I_a \cup I_b)} \times \frac{P(H_p|I_a \cup I_b)}{P(H_d|I_a \cup I_b)}.$$

#### 4. STATISTIKA V KAZENSKEM PRAVU

Statistika se lahko nanaša tudi na različne vrste meritev. Znani primeri v kazenskih postopkih so analize kemične sestave sumljivih snovi (kot so droge ali strupi) in meritve elementne sestave steklenih delcev. Čeprav se tovrstne forenzične statistike rutinsko vključujejo v dokaze, predložene v kazenskih postopkih, pa bi lahko vsaka vrsta statističnih informacij načeloma postala predmet spornega vprašanja v kazenskem postopku. Te meritve se včasih splošno imenujejo "šporenljivke", saj se razlikujejo od predmeta do predmeta (npr. različna kemična vsebnost narkotičnih tablet, različna elementna sestava steklenih odlomkov itd.)

Ocenjevanje verjetnosti je odvisno od dveh dejavnikov: dogodka E, katerega verjetnost se obravnava, in informacij I, ki so ocenjevalcu na voljo, ko se obravnava verjetnost E. Rezultat take ocene je verjetnost, da se bo dogodek E zgodil, če so znane I. Vse verjetnosti so pogojene z določenimi informacijami. Dogodek E je lahko sporen dogodek v preteklosti.

Statistične metode se lahko uporabljajo tudi za razlago podatkov in vrednotenje dokazov, poleg tega pa se lahko prispevajo tudi dokazi v obliki statističnih podatkov. Statistični podatki se ne uporabljajo le kot podatki z dokaznim pomenom za razjasnitev spornih dejstev, ki bi jih bilo mogoče predložiti sodišču kot izvedenski dokaz, temveč tudi kot podlaga za nadaljnje sklepanje, katerega ustreznost se lahko oceni z uporabo statističnih metod in teorije verjetnosti.

Koristno je razlikovati med dvema vrstama vzorcev, ki sta značilni za ocenjevanje znanstvenih dokazov v kazenskih postopkih. Žal ni standardne ali dogovorjene terminologije za izražanje ustreznega razlikovanja med (i) vzorci znanega izvora in

(ii) vzorci neznanega izvora v zvezi z vprašanjem v zadevi. Vzorec neznanega izvora je lahko opisan kot odvzet vzorec ali sporni vzorec, medtem ko so vzorci znanega izvora pogosto opisani kot kontrolni vzorec ali referenčni vzorec. Kontrolni/referenčni vzorec je lahko vzet s kraja kaznivega dejanja, od žrtve ali osumljenca. Nasprotno pa lahko izterjani/izpraševani vzorec prav tako izvira iz katerega koli od teh virov.

Statistične metode se lahko uporabljajo tudi za pridobivanje novih podatkov, ki se uporabljajo v forenziki. Najprej je potrebno opredeliti forenzični problem, nato se določi, katere informacije so pomembne, to pa preiskovalcu omogoča, da oceni, kako bi lahko ustvaril zanesljiv vzorec, da bi pridobil nove podatke za utemeljene sklepne ugotovitve.

## 5. OSNOVNI KONCEPTI VERJETNOSTNEGA SKLEPANJA IN DOKAZOV

Da bi lahko razlagali in ocenjevali statistične dokaze ter ocenili ustreznost morebitnih verjetnostnih sklepov, ki naj bi jih dokazi podpirali, potrebujemo naslednje pojme:

- (absolutne in relativne) frekvence(angl. (absolute and relative) frequencies);
- razmerje verjetnosti dokazov(angl. likelihood of the evidence);
- razmerje verjetnosti(angl. the likelihood ratio);
- predhodne verjetnosti(angl. prior probabilities);
- posteriorne verjetnosti(angl. posterior probabilities);
- Bayesova teorija(angl. Bayes' Theorem);
- neodvisnost(angl. independence).

### 5.1. Frekvence(absolutne in relativne)(angl. (absolute and relative) frequencies).

## 6. NAPAKE

6.1. **Verjetnostna napaka vira(angl. Source probability error).** Kadar pride do nezakonitih prenosov pogojnika glede na propozicije na ravni vira, je to bolj strokovno znano kot napaka verjetnosti vira.

## 7. KAKO SE STATISTIKA UPORABLJA V KAZENSKEM PRAVU

Izhodišče so nekateri statistični podatki, npr. kolikokrat se profil DNK pojavi v podatkovni zbirki; zapisi o prodaji, potovanjih ali izmenah v bolnišnicah. Nato se iz podatkov na podlagi statističnega modela izpelje statistična ali verjetnostna ocena, npr. ocena pogostosti profila DNK v populaciji se izpelje na podlagi populacijskega modela, ali pa se na podlagi določenih predpostavk izpelje verjetnost nekega dogodka. Nazadnje se verjetnostne in statistične ocene uporabijo za sklepanje o krivdi ali nedolžnosti.