

Intelligens orvosi műszerek (vimia023)

Az elmúlt évek tapasztalatai alapján ugyanannak a dolognak az alábbi *kétféle megadási lehetősége* egyeseket megzavar; megpróbálom tisztába tenni.

Arról akarunk információt adni: valamilyen vizsgált csoportnak a mért (jelen esetben z -vel jelölt) paramétere hogyan oszlik meg a csoportban. *Tehát ugyanarról van szó mindkét esetben, csak az információ megadásának formája más.*

Olyan ez, mintha egyrészt azt mondanám:

„A magyar állampolgárok 7%-ának a magassága kevesebb mint 160 cm, 12 %-ának 160 és 170 cm közé esik, 27%-ának pedig 170 és 180 cm közé.”

Másrészt pontosan ugyanezt mondom, de másképp fogalmazom meg, ha a következőt állítom:

„A magyar állampolgárok 7%-ának a magassága 0 és 160 cm közé esik, 19%-ának (=7+12) 0 és 170 cm közé, 46%-ának (=7+12+27) pedig 0 és 180 cm közé.”

Tehát az alábbi két táblázat minden szempontból azonos információt hordoz:

j	1	2	3	4	5
$[z_{min}, z_{max})$	$[-7,10)$	$[-7,23)$	$[-7,40)$	$[-7,67)$	$[-7,83]$
Betegek, akiknek $[z_{min}, z_{max})$ – ba esik a mért z paramétere	23	91	260	485	827
Egészségesek, akiknek $[0, z_{max})$ – ba esik a mért z paramétere	874	1412	1723	1827	1869

j	1	2	3	4	5
$[z_{jmin}, z_{jmax})$	$[-7,10)$	$[10,23)$	$[23,40)$	$[40,67)$	$[67,83]$
Egészségesek, akiknek $[z_{jmin}, z_{jmax})$ – ba esik a mért z paramétere	874	538	311	104	42
Betegek, akiknek $[z_{jmin}, z_{jmax})$ – ba esik a mért z paramétere	23	68	169	225	342

Megjegyzés: 1.) A tisztesség kedvéért az $[A,B)$ stb. jelöléseket alkalmaztam az intervallumok megadásánál, ami azt jelenti, hogy az A értéket – amely egyben határa is a szóban forgó intervallumnak – felveheti ebben az intervallumban a z , a B viszont nem része, csak határa az intervallumnak. Ennek gyakorlati jelentősége az ilyen esetekben általában minimális, de így korrekt és egyértelmű.

2.) Éberségi ellenőrzés kedvéért: az első táblázat első sora tartalmazza a betegekre vonatkozó adatokat, a második sora az egészségesekre vonatkozókat, míg a második táblázatnál éppen fordított a sorrend. Bocs! ☺

Nézzük meg ezek alapján, hogy ha egy eljárásnál pl. azt javasoljuk, hogy $z < 23$ esetén tekintsük egészségesnek a vizsgált személyt, míg $z \geq 23$ esetén betegnek, akkor hogyan alakul erre az eljárásra a szenzitivitás és specificitás értéke.

Bármelyik táblázatot is nézzük, azt tapasztaljuk, hogy a vizsgált – $1869+827=2696$ személyből álló – csoportban 1412 valóban egészségesnek lesz kisebb a mért paramétere a limitnél – tehát őket helyesen egészségesnek fogjuk tekinteni ($TN=1412$), míg 91 betegnek lesz kisebb a limitnél a mért paramétere – tehát őket tévesen egészségesnek fogjuk tekinteni ($FN=91$).

Az összes egészséges a vizsgálat során 1869 ember volt, az összes beteg 827. Így $FP=$ összes negatív – $TN=1869-1412=457$, amit úgy is megkaphattunk volna, hogy összegezzük azon csoportokat, akiknek a limittel egyenlő vagy annál nagyobb a mért paramétere: $311+104+42=457$. Hasonló módon a valós pozitív detektálások száma: $TP=$ összes beteg – $FN=827-91=736$, vagy a limittel egyenlő vagy a feletti értékkel rendelkező csoportokból: $169+225+342=736$.

Innen a keresett paraméterek:

szenzitivitás = $TP/(TP+FN) = 736/827 = 0,89$ (89%)
specificitás = $TN/(TN+FP) = 1412/1869 = 0,755$ (75,5%)