p-vrednost testa

Pomenuli smo da testiranje hipoteza možemo izvršiti nalaženjem kritične oblasti W, koja odgovara vrednostima test statistike T koje idu u prilog alternativnoj hipotezi H_1 i koja je takva da važi:

$$P\{T \in W | H_0 \text{ je tačna}\} = \alpha,$$

gde je α unapred zadat nivo značajnosti testa.

Dotakli smo se i pristupa preko p-vrednosti testa, pa ćemo sada razjasniti koja motivacija stoji iza tog pojma. Formalno, p-vrednost testa se definiše na sledeći način:

def: p-vrednost testa je verovatnoća da test statistika T, pod uslovom da je H_0 tačna, uzme bar utoliko ekstremnu vrednost kao što je uzela na datom uzorku. Pod *ekstremnim* vrednostima mislimo na one vrednosti koje idu u prilog H_1.

Rekli smo da nam mala p-vrednost sugeriše da bi trebalo odbaciti H_0, jer ona znači da je realizovana vrednost test statistike T toliko neočekivana za nultu hipotezu da je mala verovatnoća da, pod uslovom da je H_0 zaista tačna, T uzme takvu ili još ekstremniju vrednost. Da objasnimo ovo sada i na praktičnom primeru:

Dogovorili smo se sa prijateljem da se sastanemo na jednom mestu u 18:00 h, a on je stigao u 18:45 h. Zašto bismo mogli da poverujemo da on nije krenuo na vreme ka dogovorenom mestu?

1. Da je zaista krenuo na vreme, stigao bi bar do 18:15 h (možda su se na putu desile neke nepredviđene okolnosti poput zastoja u saobraćaju, pa ćemo tolerisati ovih 15 minuta kašnjenja);

2. Da je zaista krenuo na vreme, ne bi baš ovoliko ili više zakasnio.

Dakle, prvi razlog zbog kog sumnjamo da naš prijatelj nije krenuo na vreme je taj što je njegovo kašnjenje premašilo određenu granicu, a drugi razlog je taj što je njegovo kašnjenje previše ekstremno da bismo poverovali da je zaista krenuo na vreme.

Prednosti testiranja hipoteza preko p -vrednosti

Ovaj pristup takođe zahteva poznavanje raspodele test statistike pod uslovom da je H_0 tačna, ali ne zahteva konstruisanje kritične oblasti koja, kao što smo videli, zavisi od nivoa značajnosti α. Posmatrajmo sledeći primer da vidimo i zašto:

primer 1. Neka je dat prost slučajan uzorak $(X_1, X_2, \ldots, X_n)$ iz $\mathcal{N}(m, \sigma^2)$ raspodele, gde je σ^2 poznato. Sa nivoom značajnosti α testirati:

$$H_0 : m = m_0 \quad \text{vs} \quad H_1 : m < m_0$$

$a)$ preko kritične oblasti;
$b)$ preko p-vrednosti testa.

Rešenje: $a)$ Rekli smo da vrednosti statistike $\overline{X}_n$ manje od m_0 idu u prilog H_1, jer je $\overline{X}_n$ nepristrasna, a i postojana ocena parametra m, pa očekujemo da njena vrednost bude bliska pravoj vrednosti tog parametra. Dakle, što je $m_0 - \overline{x}_n$ veće, odnosno $\overline{x}_n - m_0$ manje, to sto više sigurni u H_1. Hajde zato da kritičnu oblast potražimo u sledećem obliku:

$$W = \{\overline{X}_n - m_0 < c_1\},$$

odnosno:

$$W = \{\frac{\overline{X}_n - m_o}{\sigma}\sqrt{n} < c\}$$

Ovu transformaciju kritične oblasti W smo izvršili jer će nam biti lakše da odredimo c ako za test statistiku uzmemo $T = \frac{\overline{X}_n - m_o}{\sigma}\sqrt{n}$, jer znamo njenu raspodelu u slučaju da važi H_0.

Dakle, c treba da bude takvo da je:

$$P\{T \in W | H_0 \text{ je tačna}\} = \alpha \iff P\{T < c | H_0 \text{ je tačna}\} = \alpha$$

$$\iff P\{Z < c\} = \alpha, \quad Z \sim \mathcal{N}(0,1),$$

jer T u slučaju da važi H_0 ima $\mathcal{N}(0,1)$ raspodelu, pa zaključujemo da je $c = \phi^{-1}(\alpha)$. Pošto smo našli T i W, možemo izvršiti testiranje na osnovu datih vrednosti uzorka.

b) Označimo sa t realizovanu vrednost statistike T na datom uzorku. Prateći definiciju p-vrednosti zaključujemo da je ona jednaka:

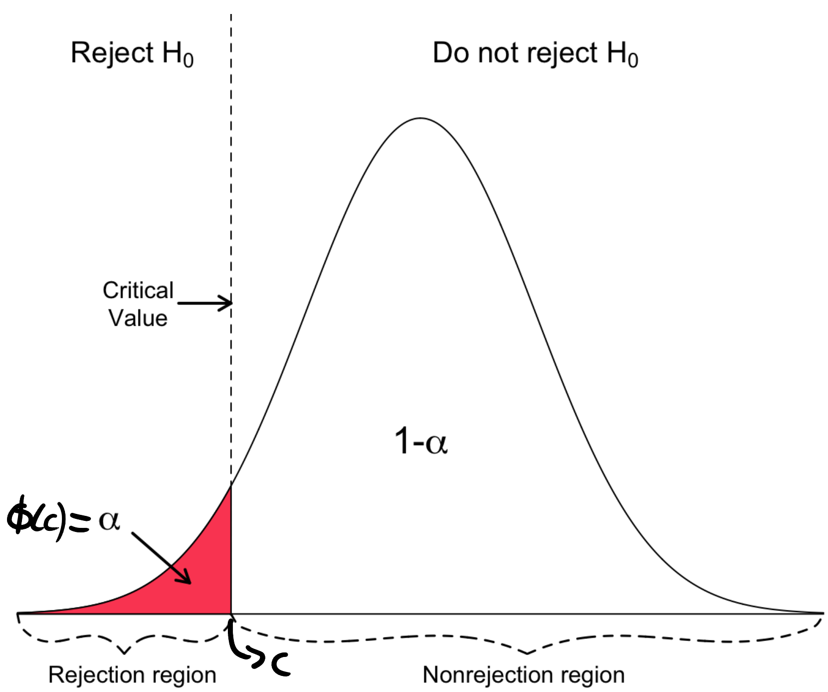
$$p = P\{T < t | H_0 \text{ je tačna}\},$$

jer su sve vrednosti T manje od t još ekstremnije za H_0. Kako T u slučaju da važi H_0 ima $\mathcal{N}(0,1)$ raspodelu, sledi:

$$p = \phi(t)$$

Preostaje nam da uporedimo p i α i da nultu hipotezu odbacimo ako je $p < \alpha$, odnosno ne odbacimo ako je $p > \alpha$.

Zašto je ovaj pristup ekvivalentan prvom? Pogledajmo sledeću sliku:



Slika: Grafik gustine test statistike pod nultom hipotezom. Možemo primetiti da $p < \alpha$ znači da je t upalo u kritičnu oblast, odnosno da je $t < c$, pa se H_0 odbacuje, a da $p > \alpha$ znači da t nije upalo u kritičnu oblast, tj. da važi $t > c$, pa se H_0 ne odbacuje.

Dakle, u pristupu preko p-vrednosti **ne moramo** konstruisati W da bismo izvršili testiranje sa odgovarajućim nivoom značajnosti testa, što nam omogućava da ovako mnogo brže i lakše izvršimo testiranje za različite nivoe α, jer je dovoljno da p-vrednost samo uporedimo sa svakim od nivoa, dok bi pristup preko kritične oblasti zahtevao nalaženje broja c za svako dato α!