

Košarkaški primjer

Lokalni košarkaš tvrdi da pogađa 80% svojih slobodnih bacanja. Kada ga zamolite da to pokaže, on izvede 20 slobodnih bacanja i pogodi samo 8. Na osnovu ovog rezultata, zaključujete da je njegova tvrdnja malo vjerovatna. Ovo je primjer statističkog testa značajnosti na nivou igrališta.

Statističko zaključivanje

Statističko zaključivanje koristi podatke iz uzorka za izvođenje zaključaka o populaciji. Kada napustimo igralište, statistički testovi se bave tvrdnjama o populaciji. Oni pitaju da li podaci uzorka daju dobar dokaz protiv tvrdnje. Statistički test kaže: "Kada bismo uzeli mnogo uzoraka i tvrdnja je tačna, rijetko bismo dobili ovakav rezultat." Da biste dobili numeričku mjeru koliko je dokaz uzorka jak, zamijenite nejasan izraz "rijetko" vjerovatnoćom.

Eksperiment sa kafom

Ljudi sa ukusom bi trebali preferirati svježu kafu od instant varijante. Ali, možda je mnogima koji piju kafu samo potreban kofeinski preparat. Skeptik tvrdi da oni koji piju kafu ne mogu razlikovati. Uradimo eksperiment da testiramo ovu tvrdnju. Svaki od 50 ispitanika proba dvije neoznačene šolje kafe i kaže koja mu je draža. Jedna šoljica u svakom paru sadrži instant kafu; druga, svježe kuhana kafa.

Rezultati eksperimenta

Statistički podatak koji bilježi rezultat našeg eksperimenta je udio uzorka \hat{p} koji kažu da im se više sviđa svježe kuhana kafa. Otkrili smo da 36 od naših 50 ispitanika bira svježu kafu. To je $\hat{p} = \frac{36}{50} = 0.72 = 72\%$. Da bismo naglasili poentu, uporedimo naš ishod $\hat{p} = 0.72$ sa drugim mogućim rezultatom. Ako samo 28 od 50 ispitanika voli svježu kafu više od instant kafe, proporcija uzorka je $\hat{p} = \frac{28}{50} = 0.56 = 56\%$.

Jačina dokaza

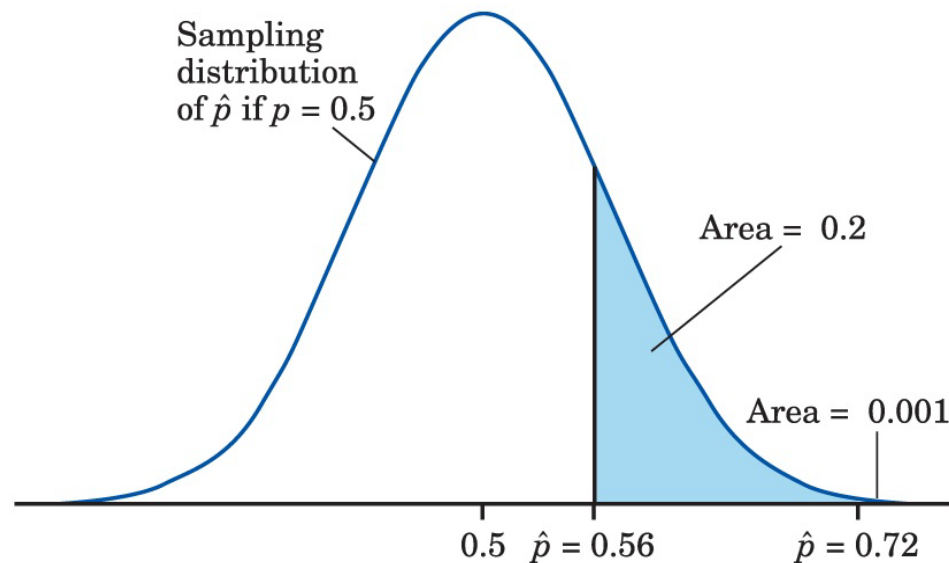
Sigurno je 72% jači dokaz protiv tvrdnje skeptika od 56%. Ali koliko jači?

Tvrdnja skeptika

Skeptik tvrdi da oni koji piju kafu ne mogu razlikovati svježu od instant, pa će samo polovina odabrati svježu kafu. Odnosno, on tvrdi da je udio stanovništva p samo 0.5.

Distribucija uzorkovanja

Da je tvrdnja $p = 0.5$ tačna i da smo testirali mnoge nasumične uzorke od 50 ljudi koji piju kafu, udio uzorka \hat{p} bi varirao od uzorka do uzorka prema (približno) normalnoj distribuciji (sa 0.5 kao srednjom vrijednošću).



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company

Nulta i alternativna hipoteza

- **Nulta hipoteza (H0):** Tvrdnja koja se testira u statističkom testu. Obično je to izjava "bez efekta" ili "nema razlike".
- **Alternativna hipoteza (Ha):** Tvrdnja za koju se nadamo ili sumnjamo da je istinita umjesto H0.

p-vrijednost

- **p-vrijednost:** Vjerovatnoća, izračunata pod pretpostavkom da je H0 istinita, da bi ishod uzorka bio jednako ekstreman ili ekstremniji od stvarno promatranog ishoda. Što je p-vrijednost manja, to su podaci jači dokaz protiv H0.

Test značajnosti

- **Test značajnosti:** Traži dokaze protiv nulte hipoteze a u korist alternativne hipoteze. Dokazi su jaki ako bi se ishod koji promatramo rijetko dogodio ako je nulta hipoteza tačna, ali je vjerojatniji ako je alternativna hipoteza istinita.

Primjer sa studentima

Prema izvještaju Centra za obrazovanje i radnu snagu Univerziteta Georgetown iz 2015. godine, 70% studenata u Sjedinjenim Državama ima posao na puno ili skraćeno radno vrijeme dok su upisani na fakultet ($p = 0.7$). Administrator sa lokalnog koledža dovodi u pitanje tačnost ove tvrdnje i vjeruje da se pravi udio studenata na njenom koledžu razlikuje od 0.7.

Eksperiment

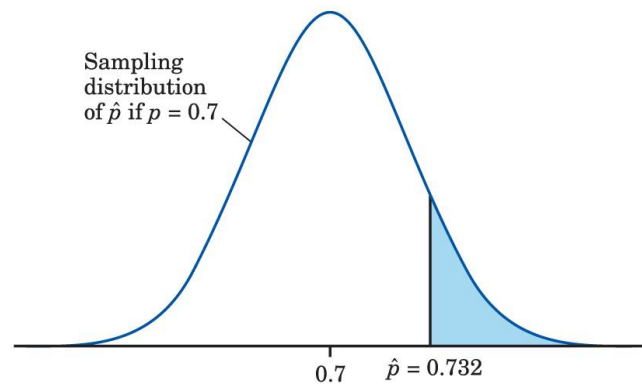
- **Uzorak:** 325 studenata
- **Rezultat:** 238 od 325 studenata ima posao ($p^{\wedge} = 238/325 = 0.732$)

Hipoteze

- **Nulta hipoteza (H0):** $p = 0.7$
- **Alternativna hipoteza (Ha):** $p \neq 0.7$

Distribucija uzorkovanja

Ako je nulta hipoteza tačna, udio u uzorku studenata koji rade puno ili skraćeno radno vrijeme ima približno normalnu distribuciju sa srednjom vrijednošću $p = 0.7$ i standardnom devijacijom $\sqrt{\frac{0.7 \cdot 0.3}{325}} = 0.02542$.



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company

p-vrijednost

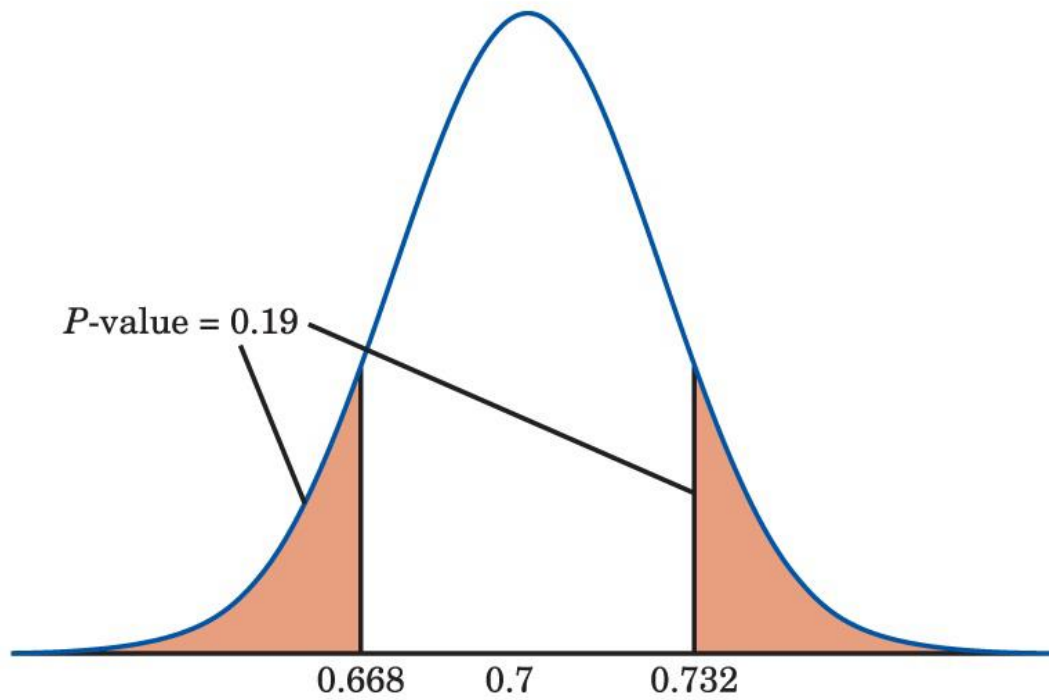
- **p-vrijednost:** Vjerovatnoća da je opaženi \hat{p} udaljen od 0.7 u oba smjera koliko i opaženi $\hat{p} = 0.732$.
Prema izračunu, p-vrijednost je 0.19.

Zaključak

Ako je pravi udio populacije 0.7, vjerovatnoća da ćemo dobiti proporciju uzorka ovako daleko ili dalje od 0.7 je 0.19. Stoga ne možemo odbaciti tvrdnju da 70% studenata radi puno ili skraćeno radno vrijeme dok pohađa fakultet.

Jednostrana i dvostrana alternativa

- **Jednostrana alternativa ($H_a: p > 0.5$):** Efekat za koji tražimo dokaze kaže da je udio populacije veći od jedne polovine.
- **Dvostrana alternativa ($H_a: p \neq 0.7$):** Pitamo da li se udio populacije p razlikuje od 0.7. Da li je alternativa jednostrana ili dvostrana određuje da li su rezultati uzorka ekstremni u jednom ili u oba smjera.



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company

Nivo značajnosti (α)

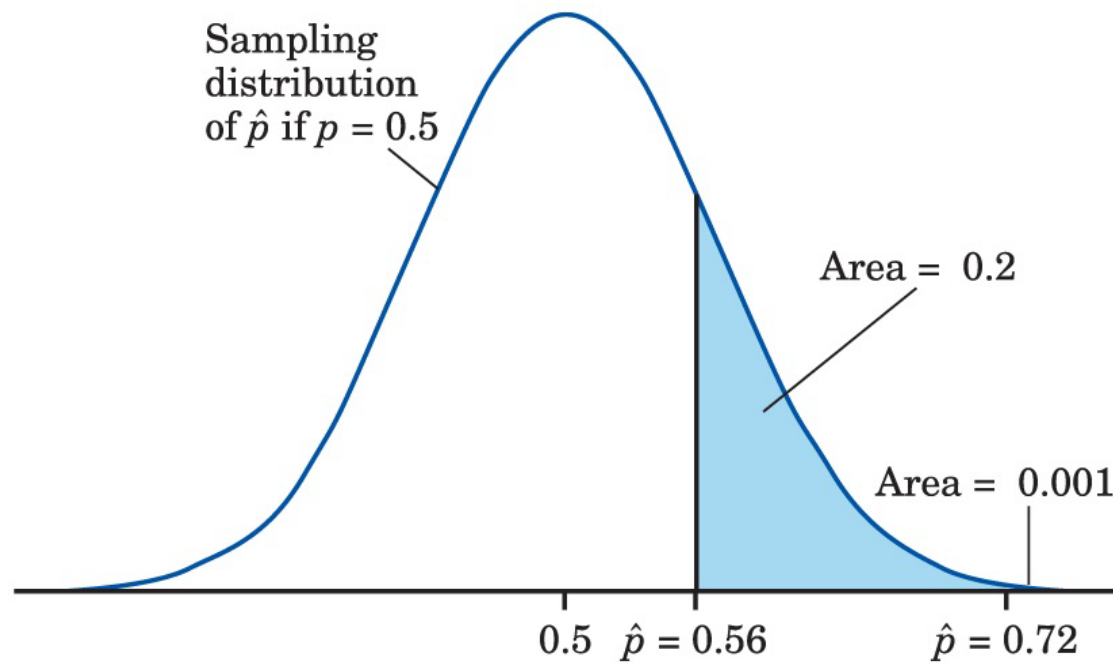
- Nivo značajnosti (α): Ako je p-vrijednost manja ili jednaka α , kažemo da su podaci statistički značajni na nivou α .
- Odabir α : Prije prikupljanja podataka, odlučujemo koliko mala p-vrijednost treba biti da bismo odbacili nultu hipotezu (H_0). Uobičajeni nivoi značajnosti su $\alpha = 0.05$ (umjeren dokaz) i $\alpha = 0.01$ (snažan dokaz).

Značajnost u statistici

- Statistički značajno: Ne znači "važno", već "nije vjerovatno da će se dogoditi slučajno".
- P-vrijednost: Informativnija od α jer omogućava procjenu značaja na bilo kojem nivou. Na primjer, $p = 0.03$ je značajno na nivou $\alpha = 0.05$, ali nije na nivou $\alpha = 0.01$.

Primjer sa kafom

- Hipoteze: $H_0: p = 0.5$, $H_a: p > 0.5$
- Uzorak: 50 ljudi, 36 preferira svježu kafu ($p^{\wedge} = 0.72$)
- Z-score: $(0.72 - 0.5) / 0.0707 = 3.1$
- P-vrijednost: 0.001 (vrlo jak dokaz protiv H_0)



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company

Primjer sa trudnoćom

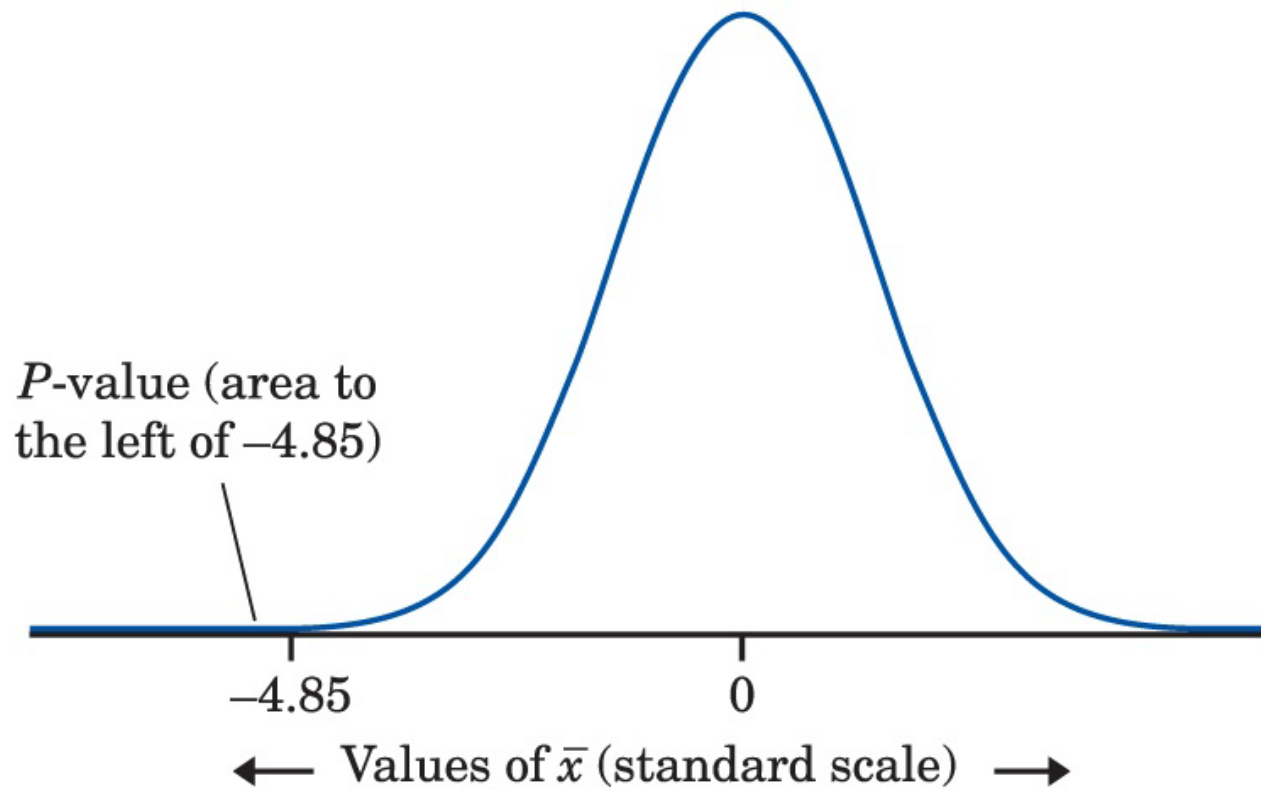
- Hipoteze: $H_0: \mu = 280$, $H_a: \mu < 280$
- Uzorak: 95 žena, srednja dužina trudnoće $\bar{x} = 275$ dana, $s = 10$ dana
- Z-score: $(275 - 280) / (10 / \sqrt{95}) = -4.85$
- P-vrijednost: Manja od 0.0003 (jak dokaz protiv H_0)

Primjer sa novčićem

- Hipoteze: $H_0: p = 0.5$, $H_a: p \neq 0.5$
- Uzorak: Grof Buffon bacio novčić 4040 puta, dobio 2048 glava ($p^{\wedge} = 0.507$)
- P-vrijednost: 0.37 za $n = 4040$ (nije dokaz protiv H_0), 0.000009 za $n = 100,000$ (jak dokaz protiv H_0)

Greške tipa I i II

- Greška tipa I: Odbijamo H_0 kada je H_0 zapravo istinit.
- Greška tipa II: Prihvatamo H_0 kada je H_a zapravo istinita.



Moore/Notz, *Statistics: Concepts and Controversies*, 10e, © 2020
W. H. Freeman and Company