

Računanje verovatnoća i rad sa funkcijama u vezi sa distribucijama verovatnoća u programu SPSS: rešenja zadataka sa vežbi

Zadatke uradila Violeta Mihajlovska, studentkinja psihologije, broj indeksa PS160065

Zadatak 1.

Otvoriti bilo koji fajl sa podacima.

- Izračunati (korišćenjem funkcije CDF.NORMAL) kolika je verovatnoća da osobe u populaciji imaju količnik inteligencije veći od 150, ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i varijansom 225. Varijablu u koju će se upisati rezultat nazvati IQpr150;
- Izračunati (korišćenjem funkcije CDF.NORMAL) kolika je verovatnoća da osobe u populaciji imaju količnik inteligencije između 85 i 115, ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i varijansom 225. Varijablu u koju će se upisati rezultat nazvati IQ85do115;

Rešenje:

- Da bismo izračunali verovatnoću da neka osoba u populaciji ima IQ veći od 150, u kojoj je inteligencija normalno distribuirana sa aritmetičkom sredinom 100 i varijansom 225, tj. standardnom devijacijom 15 (pošto je standardna devijacija jednaka korenu varijanse), koristićemo komandu CDF.NORMAL:

Biramo **Transform>Compute Variable**

(U spisku *Function group* pronaći ćemo *CDF & Noncentral CDF* (CDF je akronim od Cumulative Distribution Function), a u *Functions and Special Variables* ćemo kliknuti **Cdf.Normal** i potom ćemo kliknuti na strelicu nagore kako bismo funkciju prebacili u polje *Numeric Expression*).

U polje *Target Variable* upisaćemo IQpr150, a u polju *Numeric Expression* ćemo dopuniti izraz tako da bude $1 - \text{CDF.NORMAL}(150,100,15)$. Naime, funkcija CDF.NORMAL (150,100,15) daje verovatnoću da IQ uzme neku vrednost manju ili jednaku 150, a nama treba verovatnoća da IQ uzme neku vrednost veću od 150. Budući da je celokupna verovatnoća pod normalnom krivom jednaka 1, oduzimanjem od 1 verovatnoće koju daje funkcija CDF.NORMAL(150,100,15) dobijamo željenu verovatnoću. (Klikom na dugme **OK** dobija se verovatnoća da neka osoba ima količnik inteligencije preko 150, koja se pojavljuje kao vrednost na novoj "varijabli" IQpr150.

Verovatnoća koja se ovim postupkom dobija, zaokružena na 5 decimala, iznosi 0.00043.
Dakle, 150 je zapravo vrednost kvantila 0.99957 u ovoj raspodeli.

- Za računanje verovatnoće da nečiji IQ u populaciji u kojoj je inteligencija normalno distribuirana sa aritmetičkom sredinom 100 i varijansom 225 (standardnom devijacijom 15) uzme vrednost između 85 i 115 koristimo isti postupak sa određenim izmenama:

Transform>Compute Variable

U polje *Target Variable* unosimo IQ85do115, a u polje *Numeric Expression* $CDF.NORMAL(115,100,15) - CDF.NORMAL(85,100,15)$ i klikom na dugme OK dobijamo željeni rezultat kao vrednost na “varijabli” IQ85do115, koji zaokružen na 3 decimale iznosi 0.683.

Budući da je IQ od 115 za jednu standardnu devijaciju iznad aritmetičke sredine a da je IQ od 85 za jednu standardnu devijaciju ispod aritmetičke sredine, možemo reći da je verovatnoća da varijabla koja ima normalnu raspodelu uzme neku od vrednosti između tačaka koje su udaljene na svaku stranu za po jednu standardnu devijaciju od aritmetičke sredine jednaka 68.3%.

Zadatak 2.

Fajl isti kao za zadatak 1.

- Ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i standardnom devijacijom 15 izračunati (korišćenjem funkcije IDF.NORMAL) za koju vrednost količnika inteligencije važi sledeće: verovatnoća da osobe u populaciji imaju količnik inteligencije veći od te vrednosti jednaka je 0.005. Varijablu u koju će se upisati rezultat nazvati KojiIQ1;
- Ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i standardnom devijacijom 15 izračunati (korišćenjem funkcije IDF.NORMAL) za koju vrednost količnika inteligencije važi sledeće: verovatnoća da osobe u populaciji imaju količnik inteligencije manji od te vrednosti jednaka je 0.005. Varijablu u koju će se upisati rezultat nazvati KojiIQ2;
- Ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i standardnom devijacijom 15 izračunati (korišćenjem funkcije IDF.NORMAL) za koju vrednost standardizovanog količnika inteligencije (z-vrednost) važi sledeće: verovatnoća da osobe u populaciji imaju standardizovani količnik inteligencije veći od te vrednosti jednaka je 0.005. Varijablu u koju će se upisati rezultat nazvati Kojazvr1;
- Ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i standardnom devijacijom 15 izračunati (korišćenjem funkcije IDF.NORMAL) za koju standardizovanu vrednost količnika inteligencije (z-vrednost) važi sledeće: verovatnoća da osobe u populaciji imaju standardizovani količnik inteligencije manji od te vrednosti jednaka je 0.005. Varijablu u koju će se upisati rezultat nazvati Kojazvr2;

Rešenje:

- Da bismo došli do vrednosti za koju važi da je verovatnoća da osobe u populaciji imaju IQ iznad te vrednosti jednaka 0.005, tj. da bismo izračunali vrednost kvantila 0.995, koristimo komandu **IDF.NORMAL**.

Transform>Compute Variable

(U *Function group* ćemo pronaći **Inverse DF** (skraćenica od Inverse Distribution Function), a u *Functions and Special Variables* ćemo kliknuti na **IDF.NORMAL** i potom na strelicu nagore). *Target Variable* će biti KojiIQ1, a u *Numeric Expression* ćemo uneti IDF.NORMAL(0.995,100,15). Unosimo verovatnoću 0.995 jer komanda IDF.NORMAL zahteva da unesemo verovatnoću da se ispod vrednosti koju tražimo nađe određeni procenat slučajeva. U ovom slučaju za vrednost koju tražimo važi da je verovatnoća da se osobe iz populacije nađu ispod nje jednaka 0.995. Dakle, tražimo vrednost kvantila 0.995. Klikom na OK dobijamo rezultat koji iznosi 138.64 (zaokružen na 2 decimale).

- U ovom slučaju tražimo vrednost za koju važi da je verovatnoća da nečiji IQ u populaciji sa normalno distribuiranom inteligencijom, čija je aritmetička sredina 100 a standardna devijacija 15, bude ispod te vrednosti jednaka 0,005. Dakle, tražimo kvantil 0.005. U ovom slučaju koristimo potpuno isti postupak kao u prethodnom primeru, s tim da nam je *Target Variable* KojiIQ2 i u polje *Numeric Expression* unosimo IDF.NORMAL(0.005, 100, 15). Klikom na OK dobijamo traženu vrednost, a ona je 61.36 (zaokružena na 2 decimale).

- Da bismo došli do standardizovane IQ vrednosti (z-vrednosti) za koju važi da je verovatnoća da osobe u populaciji imaju IQ iznad te vrednosti jednaka 0.005, koristimo komandu IDF.NORMAL:

Transform>Compute Variable

(U *Function group* ćemo pronaći **Inverse DF**, a u *Functions and Special Variables* ćemo kliknuti na IDF.NORMAL pa na strelicu nagore). **Prema jednoj teoremi iz teorije verovatnoće, ako varijabla ima normalnu raspodelu, svaka linearna transformacija te varijable će takođe imati normalnu raspodelu. Prema tome, budući da IQ ima normalnu raspodelu sa parametrima 100 i 15, standardizovani IQ će imati standardizovanu normalnu raspodelu, tj. raspodelu sa parametrima 0 i 1.**

U polje *Target Variable* upisaćemo **Kojazvr1**, a u *Numeric Expression* ćemo uneti IDF.NORMAL(0.995,0,1). Klikom na OK dobijamo rezultat koji iznosi 2.58 (zaokružen na dve decimale).

U prvom zahtevu u ovom zadatku dobili smo kao vrednost kvantila 0.995 vrednost 138.64 na IQ skali. Sada smo kao vrednost kvantila 0.995 na z-skali dobili vrednost 2.58. Ukoliko od vrednosti 138.64 oduzmemo aritmetičku sredinu IQ skale, tj. 100 i potom dobijenu razliku podelimo standardnom devijacijom IQ skale, tj. 15, dobićemo zapravo

2.58. Prema tome, 138.64 je za 2.58 standardnih devijacija na IQ skali iznad aritmetičke sredine, a z-vrednost je na z skali takođe za 2.58 standardnih devijacija iznad aritmetičke sredine. Dakle, i 138.64 i 2.58 predstavljaju vrednosti kvantila 0.995, samo na različitim skalama, prva vrednost na IQ skali, a potonja na z-skali.

- U ovom primeru tražimo z-vrednost za koju važi da je verovatnoća da nečiji standardizovani IQ u populaciji sa normalno distribuiranom inteligencijom cija je aritmetička sredina 100, a standardna devijacija 15, uzme neku vrednost ispod te vrednosti jednaka 0.005. Koristimo potpuno isti postupak kao i u prethodnom primeru, s tim što nam je sada *Target Variable* **Kojazvr2** a u polje *Numeric Expression* unosimo `IDF.NORMAL(0.005, 0, 1)`. Klikom na OK dobijamo traženu vrednost, a ona je -2.58 (zaokružena na 2 decimale).

U drugom zahtevu u ovom zadatku dobili smo kao vrednost kvantila 0.005 vrednost 61.36 na IQ skali. Sada smo kao vrednost kvantila 0.005 na z-skali dobili vrednost -2.58. Ukoliko od vrednosti 61.36 oduzmemo aritmetičku sredinu IQ skale, tj. 100 i potom dobijenu razliku podelimo standardnom devijacijom IQ skale, tj. 15, dobićemo -2.58. Prema tome, 61.36 je za 2.58 standardnih devijacija na IQ skali ispod aritmetičke sredine, a z-vrednost -2.58 je na z skali takođe za 2.58 standardnih devijacija ispod aritmetičke sredine. Dakle, i 61.36 i 2.58 predstavljaju vrednosti kvantila 0.995, samo na različitim skalama, prva vrednost na IQ skali, a potonja na z-skali.

Dakle, verovatnoća da standardizovani IQ koji ima standardizovanu normalnu raspodelu uzme neku vrednost manju od -2.58 ili jednaku -2.58 ili veću od 2.58 ili jednaku 2.58 jednaka je $2 \cdot 0.005$, tj. 0.01.

Zadatak 3.

Fajl isti kao za zadatak 1.

- Izračunati (korišćenjem funkcije `CDF.NORMAL`) kolika je verovatnoća da osobe u populaciji imaju standardizovani količnik inteligencije veći od 1.96, ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i varijansom 225. Varijablu u koju će se upisati rezultat nazvati verz1;
- Izračunati (korišćenjem funkcije `CDF.NORMAL`) kolika je verovatnoća da osobe u populaciji imaju standardizovani količnik inteligencije manji od -1.96, ako je inteligencija normalno distribuirana u populaciji sa aritmetičkom sredinom 100 i varijansom 225. Varijablu u koju će se upisati rezultat nazvati verz2;

Rešenje:

- Za računanje verovatnoće da standardizovani IQ u populaciji sa normalno distribuiranom inteligencijom bude veći od 1.96 koristimo funkciju CDF.NORMAL:

Transform>Compute Variable

Kao target varijablu unosimo **verz1**, a u polje *Numeric Expression* unosimo 1-CDF.NORMAL(1.96,0,1). Kod standardizovanih vrednosti aritmetička sredina je jednaka nuli, a standardna devijacija jedinici. Klikom na dugme OK dobijamo rezultat 0.025 (zaokružen na 3 decimale)

- Da bismo izračunali verovatnoću da se neka standardizovana IQ vrednost nađe na površini levo od vrednosti -1.96 koristimo se istim postupkom kao u prethodnom primeru. Kao *Target variable* unosimo **verz2** a u polje *Numeric Expression* definišemo CDF.NORMAL(-1.96, 0, 1). Klikom na dugme OK dobijamo rezultat koji je isti kao u prethodnom primeru i iznosi 0.025.

Dakle, verovatnoća da standardizovani IQ koji ima standardizovanu normalnu raspodelu uzme neku vrednost jednaku -1.96 ili manju od -1.96 ili neku vrednost jednaku +1.96 ili veću od +1.96 jednaka je $2 \cdot 0.025$, tj. 0.05.

Zadatak 4.

Fajl isti kao za zadatak 1.

- Na testu znanja statistike (koji daje normalnu raspodelu rezultata sa aritmetičkom sredinom 50 i standardnom devijacijom 10 student je postigao rezultat 55, a na testu znanja fiziologije (koji daje normalnu raspodelu rezultata sa aritmetičkom sredinom 45 i standardnom devijacijom 4 student je postigao rezultat 48. Izračunati na kojem od ovih testova je student postigao rezultat koji je, relativno gledano, bolji, tj. na kojem od testova je student bolji od većeg procenta ostalih studenata koji su rešavali testove. Variable sa rezultatima nazvati zstat i zfiziol.

Rešenje:

- Da bismo izračunali na kom je testu student postigao relativno bolji rezultat pretvorićemo sirove rezultate u standardizovane vrednosti, jer ćemo tako dobiti položaj studentovih rezultata u odnosu na učinak ostalih studenata.

Standardizovanje iliti pretvaranje u tzv. z-skorove vršimo u meniju ***Transform>Compute Variable***, pomoću formule za izračunavanje z-skora $z = (x - M) / S$, gde je z standardni skor koji tražimo, x je izvorni rezultat, M aritmetička sredina izvornih rezultata, a S standardna devijacija izvornih rezultata.

Najpre izračunavamo z-skor rezultata iz statistike, a zatim iz fiziologije, i nakon izračunavanja vršimo poređenje dobijenih z-skorova.

U polje *Target Variable* upisujemo **zstat**, a u polje *Numeric Expression* $(55 - 50) / 10$. Klikom na **OK** dobijamo z-skor 0.5, što znači da je studentov rezultat na testu iz statistike bolji od proseka pola standardne devijacije.

Tražimo z-skor za rezultat iz fiziologije: u *Target variable* upisujemo **zfiziol**, a u polje *Numeric Expression* unosimo $(48 - 45) / 4$ i kliknemo **OK**. Dobijeni rezultat je 0.75, što znači da je na testu iz fiziologije studentov rezultat veći od proseka $3/4$ standardne devijacije, stoga zaključujemo da je student postigao bolji rezultat na testu iz fiziologije. Naime, procenat rezultata koji su ispod z-skora 0.75 je 77.3% i veći je od procenta rezultata ispod z-skora 0.5 (69.1%). Ukoliko bismo hteli da izračunamo razliku ova dva procenta, budući da su raspodele na oba testa normalne, koristili bismo razliku vrednosti funkcija $CDF.NORMAL(0.75,0,1)$ i $CDF.NORMAL(0.5,0,1)$. Ova razlika iznosi 0.082 te iskazana procentima iznosi 8.2%. (Inače, na osnovu ranije prikazanih procenata vidimo i da na varijabli koja ima standardizovanu normalnu raspodelu vrednost 0.75 predstavlja kvantil 0.773, a vrednost 0.5 predstavlja kvantil 0.691).

Zadatak 5.

Fajl isti kao za zadatak 1.

- Izračunati (korišćenjem funkcije $CDF.NORMAL$) kolika je verovatnoća da varijabla koja ima standardizovanu normalnu raspodelu (raspodelu sa aritmetičkom sredinom 0 i standardnom devijacijom 1) uzme neku vrednost manju od -1.96 ili veću od 1.96. Varijablu sa rezultatom nazvati p196;
- Izračunati (korišćenjem funkcije $CDF.NORMAL$) kolika je verovatnoća da varijabla koja ima standardizovanu normalnu raspodelu (raspodelu sa aritmetičkom sredinom 0 i standardnom devijacijom 1) uzme neku vrednost manju od -2.58 ili veću od 2.58. Varijablu sa rezultatom nazvati p258;

Rešenje:

- S obzirom na to da tražimo verovatnoću da varijabla sa standardizovanom normalnom distribucijom uzme vrednost manju od -1.96 ili veću od 1.96, možemo računati samo verovatnoću da zauzme vrednost manju od -1.96 i duplirati njenu vrednost, jer su površine ispod -1.96 i iznad 1.96 jednake.

Biramo **Transform>Compute Variable** i potom u polje *Target Variable* upisujemo p196, a u polje *Numeric Expression* unesemo $CDF.NORMAL(-1.96,0,1) * 2$. Klikom na **OK** dobijamo rezultat 0.05.

- S obzirom na to tražimo verovatnoću da varijabla sa standardizovanom normalnom distribucijom uzme vrednost manju od -2.58 ili veću od 2.58, možemo računati samo verovatnoću da zauzme vrednost manju od -2.58 i duplirati tako dobijenu verovatnoću, jer su površine ispod -2.58 i iznad 2.58 jednake.

Biramo **Transform>Compute Variable** i potom u polje *Target Variable* upisujemo p258, a u polje *Numeric Expression* unesemo $CDF.NORMAL(-2.58,0,1) * 2$. Klikom na **OK** dobijamo rezultat 0.01.

Dakle, verovatnoća da varijabla koja ima standardizovanu normalnu raspodelu uzme neku vrednost manju od -1.96 ili veću od +1.96 iznosi 0.05, a verovatnoća da takva varijabla uzme neku vrednost manju od -2.58 ili veću od +2.58 iznosi 0.01. Očigledno, verovatnoća da varijabla koja ima standardizovanu normalnu raspodelu uzme neku vrednost između -1.96 i +1.96 jednaka je 0.95, a verovatnoća da takva varijabla uzme neku vrednost između -2.58 i +2.58 jednaka je 0.99.