

# TRANSFORMACIJE PODATAKA

---

## *Rangiranje*

*Rangiranje* je transformacija kvantitativne (intervalne ili omjerne) varijable u ordinalnu, odnosno zamjena rezultata odgovarajućim *rangovima*.

*Rang* je redni broj entiteta utvrđen sortiranjem prema nekoj kvantitativnoj varijabli.

# TRANSFORMACIJE PODATAKA

---

## *Rangiranje*

Uobičajeno je da se najbolji rezultat zamjenjuje rangom 1. Sukladno tome rang se entiteta u *normalno skaliranim varijablama* utvrđuje silaznim sortiranjem, a u *obrnuto skaliranim varijablama* uzlaznim sortiranjem.

*Normalno skalirana varijabla* - najveća numerička vrijednost predstavlja najbolji rezultat (npr. *Skok uvis*)

*Obrnuto skalirana varijabla* - najmanja numerička vrijednost predstavlja najbolji rezultat (npr. *Sprint 100 metara*)

# TRANSFORMACIJE PODATAKA

## *Rangiranje*

*Primjer:* 10 sportaša je izmjereno testom *Skok udalj s mjesta*. Rangiranje je izvršeno sljedećim postupkom:

	SDM		SDM		RANG	
D.P.	315	silazno sortiranje	B.V.	330	B.V.	1
J.K.	260		D.P.	315	D.P.	2
B.V.	330		D.R.	295	D.R.	3
G.I.	265		T.K.	290	T.K.	4
I.M.	240		Z.B.	275	Z.B.	5
T.K.	290		D.Š.	275	D.Š.	5
D.R.	295		I.K.	270	I.K.	7
I.K.	270		Z.P.	270	Z.P.	7
S.J.	235		G.I.	265	G.I.	9
Z.P.	270		J.K.	260	J.K.	10

# MICROSOFT EXCEL

---

## *Utvrdjivanje rangova*

Utvrdjivanje ranga vrši se pomoću funkcije *Rank*. Funkcija se unosi u označeno polje matrice odabirom opcije *Function...* padajućeg izbornika *Insert*. U traku *Order* dijaloškog okvira za unos ove funkcije potrebno je upisati *1* ako se utvrđuje rang u obrnuto skaliranoj varijabli.

**Zadatak** - U datoteci *Judo.xls* utvrdite rangove judaša na temelju rezultata u varijabli *BML* (*Bacanje medicinke iz ležanja na leđima*). Iskoristite funkciju *Rank*.

# STATISTICA 7

## *Utvrđivanje rangova*

Utvrđivanje ranga vrši se putem dijaloškog okvira *Rank Order Values* koji se pokreće odabirom opcije *Rank...* padajućeg izbornika *Data*. U dijaloškom okviru potrebno je odabrati opcije *Type of ranks – regular* i *Ranks for ties – low* te za normalno skalirane varijable *Assign rank 1 to – largest value*, odnosno za obrnuto skalirane varijable *Assign rank 1 to –smallest value*.

**Zadatak** - U datoteci *Judo.sta* utvrdite rangove judaša na temelju rezultata u varijabli *BML* (*Bacanje medicinke iz ležanja na leđima*).

# TRANSFORMACIJE PODATAKA

## *Centili*

*Centil ili percentil* je relativni rang, a izračunava se kao postotni udio ranga nekog entiteta u ukupnom broju entiteta u uzorku. Centil nekog rezultata pokazuje koliko posto entiteta u uzorku ima jednak ili lošiji rezultat.

$$CENTIL = \frac{n + 1 - RANG}{n} \cdot 100$$

gdje je

- ➔ *RANG* - rang entiteta
- ➔ *n* - broj entiteta u uzorku

# TRANSFORMACIJE PODATAKA

## *Centili*

*Primjer:* 10 sportaša je izmjereno testom *Skok udalj s mjesta*. Utvrđivanje centila je izvršeno sljedećim postupkom:

	SDM		RANG		CENTIL	
D.P.	315		B.V.	1	B.V.	100
J.K.	260		D.P.	2	D.P.	90
B.V.	330		D.R.	3	D.R.	80
G.I.	265		T.K.	4	T.K.	70
I.M.	240	rangiranje	Z.B.	5	Z.B.	60
T.K.	290	→	D.Š.	5	D.Š.	60
D.R.	295		I.K.	7	I.K.	40
I.K.	270		Z.P.	7	Z.P.	40
S.J.	235		G.I.	9	G.I.	20
Z.P.	270		J.K.	10	J.K.	10

računanje centila →

# MICROSOFT EXCEL

---

## *Izračunavanje centila*

Prije izračunavanja centila potrebno je utvrditi rangove entiteta pomoću funkcije *Rank* i broj entiteta u uzorku. Nakon toga se u prazno polje matrice unosi formula za izračunavanje centila.

**Zadatak** - U datoteci *Judo.xls* utvrdite centile judaša na temelju rezultata u varijabli *BML* (*Bacanje medicinke iz ležanja na leđima*).



# STATISTICA 7

## *Izračunavanje centila*

Prije izračunavanja centila potrebno je utvrditi rangove entiteta. Nakon toga se u traku *Long name* dijaloškog okvira za formatiranje varijabli unese *Spreadsheet formula* za izračunavanje centila (npr. Ako u uzorku ima 60 entiteta, a *BML* je naziv varijable u kojoj se nalaze prethodno utvrđeni rangovi entiteta, onda *Spreadsheet formula* za izračunavanje centila treba glasiti:  $= (60 + 1 - BML) / 60 * 100$  ).

**Zadatak** - U datoteci *Judo.sta* utvrdite centile judaša na temelju rezultata u varijabli *BML* (*Bacanje medicinke iz ležanja na leđima*). Iskoristite *Spreadsheet formulu*.

# TRANSFORMACIJE PODATAKA

---

## *Standardizacija podataka (z-vrijednosti)*

*Standardizacija podataka* je transformacija originalnih rezultata u *z-vrijednosti*.

*Standardizirani rezultat ili z-vrijednost* je odstupanje rezultata entiteta od aritmetičke sredine izraženo u dijelovima standardne devijacije.

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

Standardizacija podataka provodi se pomoću formule:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}$$

gdje je

- ➔  $z_{ij}$  - standardizirani rezultat entiteta  $i$  ( $i = 1, \dots, n$ ) u varijabli  $j$  ( $j = 1, \dots, m$ )
- ➔  $x_{ij}$  - originalna vrijednost ispitanika  $i$  u varijabli  $j$
- ➔  $\bar{x}_j$  - aritmetička sredina varijable  $j$
- ➔  $\sigma_j$  - standardna devijacija varijable  $j$

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

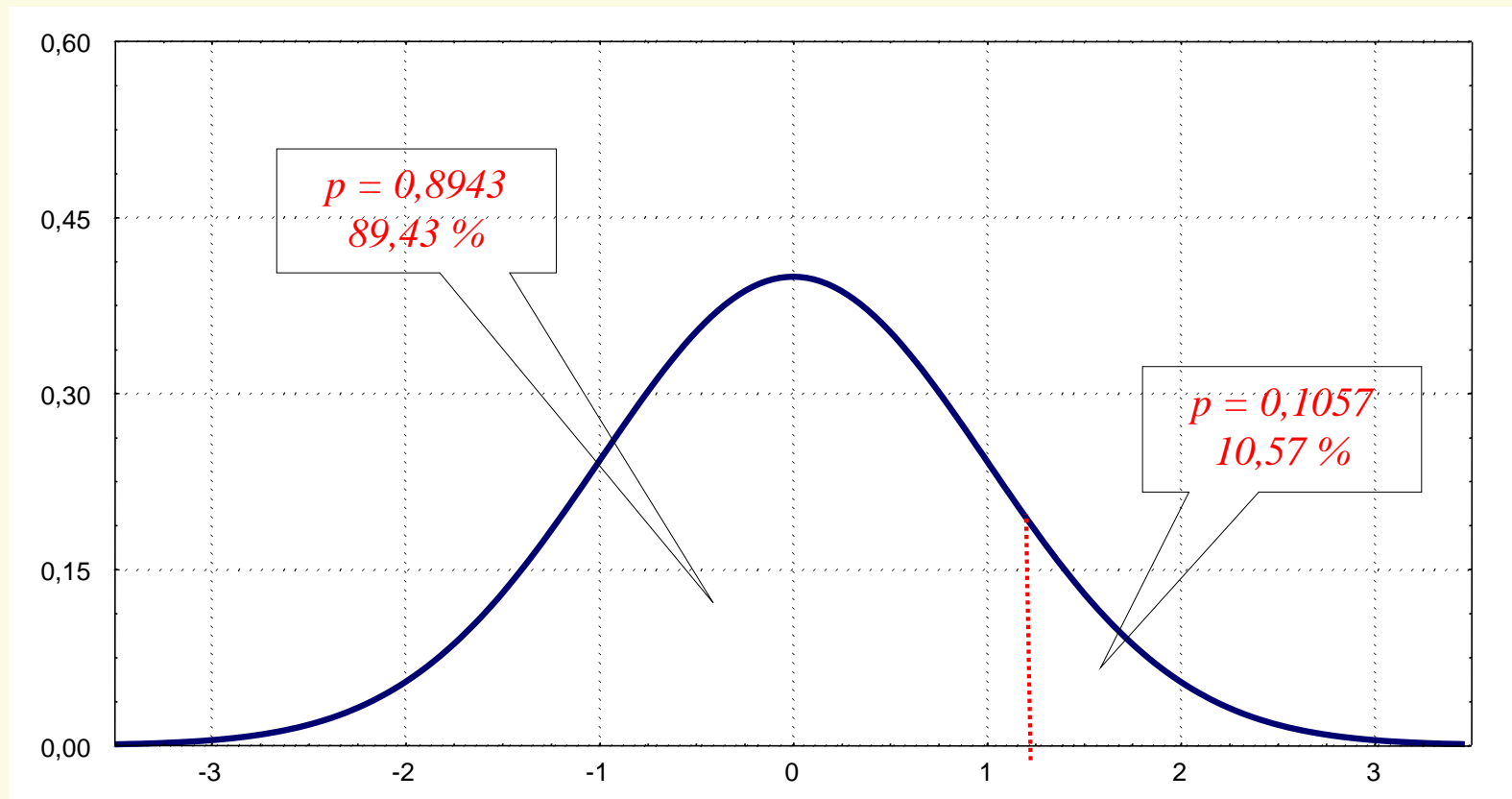
*Primjer:* 257 dječaka je izmjereno testom za procjenu eksplozivne snage *Skok udalj s mjesta*. Aritmetička sredina iznosi 215 cm, a standardna devijacija 12 cm. Učenik XY postigao je rezultat 230 cm. Potrebno je procijeniti postotak (%) i broj učenika koji postižu lošije i bolje rezultate od učenika XY.

Standardizirani rezultat izračunava se na sljedeći način:

$$z_{XY} = \frac{230 - 215}{12} = \frac{15}{12} = 1,25$$

# TRANSFORMACIJE PODATAKA

## Standardizacija podataka (z-vrijednosti)



$$z = 1,25 \Rightarrow p = 0,1057 \Rightarrow 10,57 \% \quad z = 1,25$$

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

Na temelju procijenjene vjerojatnosti može se izračunati broj ispitanika s boljim odnosno lošijim rezultatom

$$d = p \cdot n = 0,1057 \cdot 257 = 27,16 \approx 27 \quad i \quad 257 - 27 = 230$$

jer je

$$p = \frac{d}{n}, \text{ odnosno } \% = \frac{d}{n} \cdot 100$$

gdje je

➔  $p$  proporcija ( $p=0,1057$ )

➔  $d$  dio cjeline (broj učenika s boljim rezultatom od  $z=1,25$ )

➔  $n$  cjelina (ukupan broj učenika  $n=257$ )

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

*Primjer:* Deset učenika natjecalo se u tri atletske discipline:

- skok udalj (SD),
- trčanje na 100 metara (T100m) i
- bacanje kugle (BK)

i postiglo rezultate navedene u tablici. Potrebno je utvrditi ukupan poredak ovog natjecanja.

	SD	T100M	BK
A.B.	359	13,6	561
D.F.	321	13,9	550
J.G.	346	13,7	538
K.L.	332	14,0	490
D.D.	450	12,2	518
E.D.	314	14,1	551
T.B	410	12,5	589
Z.N	425	12,3	602
R.G.	369	13,5	547
E.N.	378	13,8	510

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

*Prvi korak:* Izračunati aritmetičke sredine i standardne devijacije za svaku varijablu (disciplinu).

	SD	T100m	BK
$\bar{x}$	370,4	13,36	545,6
$\sigma$	45,66	0,73	34,21

	SD	T100M	BK
A.B.	359	13,6	561
D.F.	321	13,9	550
J.G.	346	13,7	538
K.L.	332	14,0	490
D.D.	450	12,2	518
E.D.	314	14,1	551
T.B	410	12,5	589
Z.N	425	12,3	602
R.G.	369	13,5	547
E.N.	378	13,8	510



# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

*Drugi korak:* Transformirati originalne rezultate u z-vrijednosti.

$$z_{AB,SD} = \frac{359 - 370,4}{45,66} = \frac{-11,4}{45,66} = -0,25$$

	SD	T100m	BK
$\bar{x}$	370,4	13,36	545,6
$\sigma$	45,66	0,73	34,21

	SD	T100M	BK
A.B.	-0,25	0,33	0,45
D.F.	-1,08	0,74	0,13
J.G.	-0,53	0,46	-0,22
K.L.	-0,84	0,87	-1,63
D.D.	1,74	-1,58	-0,81
E.D.	-1,24	1,01	0,16
T.B	0,87	-1,17	1,27
Z.N	1,20	-1,44	1,65
R.G.	-0,03	0,19	0,04
E.N.	0,17	0,60	-1,04

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

*Treći korak:* Prije kondenzacije rezultata varijable koje su obrnuto skalirane pomnožiti s  $-1$ , odnosno rezultatima promijeniti predznake.

	SD	T100M	BK
A.B.	-0,25	-0,33	0,45
D.F.	-1,08	-0,74	0,13
J.G.	-0,53	-0,46	-0,22
K.L.	-0,84	-0,87	-1,63
D.D.	1,74	1,58	-0,81
E.D.	-1,24	-1,01	0,16
T.B	0,87	1,17	1,27
Z.N	1,20	1,44	1,65
R.G.	-0,03	-0,19	0,04
E.N.	0,17	-0,60	-1,04

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

*Četvrti korak:* Standardizirane vrijednosti kondenzirati putem aritmetičke sredine.

$$\begin{aligned} \bar{z}_{AB} &= \frac{z_{AB,SD} + z_{AB,T100} + z_{AB,BK}}{3} = \\ &= \frac{-0,25 + (-0,33) + 0,45}{3} = -0,04 \end{aligned}$$

	SD	T100M	BK	$\bar{z}$
A.B.	-0,25	-0,33	0,45	-0,04
D.F.	-1,08	-0,74	0,13	-0,56
J.G.	-0,53	-0,46	-0,22	-0,41
K.L.	-0,84	-0,87	-1,63	-1,11
D.D.	1,74	1,58	-0,81	0,84
E.D.	-1,24	-1,01	0,16	-0,70
T.B	0,87	1,17	1,27	1,10
Z.N	1,20	1,44	1,65	1,43
R.G.	-0,03	-0,19	0,04	-0,06
E.N.	0,17	-0,60	-1,04	-0,49

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

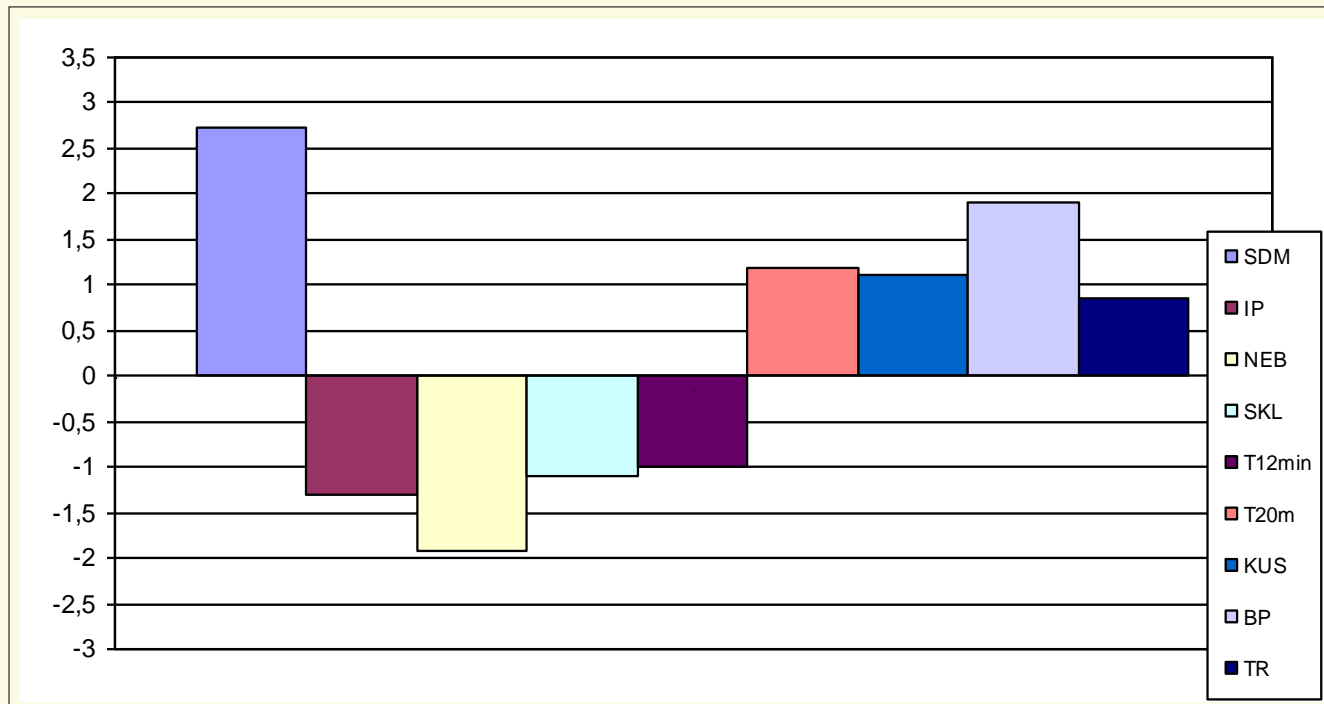
*Peti korak:* Silazno sortirati učenike po izračunatoj prosječnoj z-vrijednosti.

Rang	Učenik	$\bar{z}$
1.	Z.N	1,43
2.	T.B	1,10
3.	D.D.	0,84
4.	A.B.	-0,04
5.	R.G.	-0,06
6.	J.G.	-0,41
7.	E.N.	-0,49
8.	D.F.	-0,56
9.	E.D.	-0,70
10.	K.L.	-1,11

# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

Standardizacija podataka može se iskoristiti i u svrhu grafičkog prikazivanja rezultata entiteta u većem broju varijabli (npr. antropološki profil).



# TRANSFORMACIJE PODATAKA

## *Standardizacija podataka (z-vrijednosti)*

Standardizirane rezultate moguće je transformirati na različite načine, zavisno o potrebi. Najčešće se transformiraju u vrijednosti na sljedećim skalama:

→ skala školskih ocjena (1-5) →  $ocjena = 3 + 0,83 \cdot z$

→ skala T-skorova (0-100) →  $T - skor = 50 + 10 \cdot z$

→ skala ocjena (0-10) →  $ocjena = 5 + z$

# MICROSOFT EXCEL

## *Izračunavanje z-vrijednosti*

Izračunavanje z-vrijednosti vrši se pomoću funkcije *Standardize*. Funkcija se unosi u označeno polje matrice odabirom opcije *Function...* padajućeg izbornika *Insert*. Prethodno je potrebno izračunati aritmetičku sredinu (funkcija *Average*) i standardnu devijaciju (funkcija *Stdev*) varijable.

**Zadatak** - U datoteci *Judo.xls* izvršite standardizaciju varijabli, kondenzirajte ih aritmetičkom sredinom i utvrdite rang ispitanika prema kondenziranom rezultatu.

# STATISTICA 7

## *Izračunavanje z-vrijednosti*

Izračunavanje z-vrijednost izvodi se slijedom koraka: padajući izbornik *Edit* → *Fill/Standardize Block* → *Standardize Columns*. Kondenzacija rezultata aritmetičkom sredinom izvodi se slijedom koraka: padajući izbornik *Statistics* → *Statistics of Block Data* → *Block Rows* → *Means*. Prije oba postupka potrebno je u matrici podataka označiti željene varijable.

**Zadatak** - U datoteci *Judo.sta* izvršite standardizaciju varijabli, kondenzirajte ih aritmetičkom sredinom i utvrdite rang ispitanika prema kondenziranom rezultatu.



# STATISTICA 7

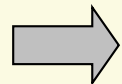
## *Utvrđivanje vjerojatnosti putem normalne distribucije*

Izračunavanje površina ispod normalne distribucije vrši se pomoću dijaloškog okvira *Probability Distribution Calculator* koji se pokreće slijedom koraka: padajući izbornik *Statistics* → *Probability Calculator* → *Distributions* → *Z (Normal)*.

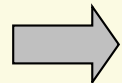
**Zadatak** - Aritmetička sredina svih službenih skokova udalj nekog atletičara iznosi 8,5 metara, a standardna devijacija 0,1 metar. Putem normalne distribucije procijenite kolika je vjerojatnost da ovaj atletičar na natjecanju izvede skok dalji od 9 metara.

# TRANSFORMACIJE PODATAKA

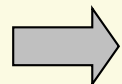
## *Literatura za pripremanje kolokvija*



Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet, str. 114-123.



Petz, B. (2002). *Osnovne statističke metode za nematematičare*. Jastrebarsko: Naklada Slap, str. 97-110.



Langer, M. (2004). *Brzi vizualni vodič Microsoft Excel 2003 za Windows*. Zagreb: Miš, str. 75-103.