



**Zdravstvena fakulteta
Smer fizioterapija –
Metode raziskovalnega dela**

31.5.2011 – Vaja 1

Podatki za analize v SPSSu

Podatki so **prirejeni in anonimizirani** na podlagi dejanskih podatkov iz diplomskega dela na Zdravstveni fakulteti (smer fizioterapija), ki je bilo nagrajeno s Prešernovo nagrado:

T.Turk: Pogostost urinske inkontinence pri ljubljanskih osnovnošolskih učiteljicah, 2009; mentorica S. Hlebš, somentor G. Vidmar

Imamo torej podatke o

- statističnih enotah: ljubljanskih osnovnošolskih učiteljicah
- spremenljivkah:

KRATKO ime	VPRAŠANJE	VREDNOSTI IN POMEN
Starost		leta
ITM	Indeks telesne mase	m ² /kg
LetPoucuje	Koliko let poučuje?	
DeloPretežno	Je vaše delo pretežno	1=sedeče ali mešano 2=stoječe
Triada	V kateri triadi pretežno poučujete	
Nosecnost	Kolikokrat ste bili noseči?	0=nikoli, 1=enkrat, 2=dva 3=več kot dvakrat
Porod	Na kak način ste rodili	0=nikoli, 1=carski rez, 2=vaginalno, 3=vaginalno car.rez
GinekoPoseg	Ste imeli kdaj kak ginekološki poseg?	0, 1
PogostaVnetjaSM	Ste imeli pogosta vnetja sečnega mehurja?	0, 1
OtrPonMocPosDalj	Ste kot otrok ponoči močili posteljo daljše obdobje?	0, 1
PogostoZaprtje	Ste pogosto zaprti (trdo blato, ni odvaj.)?	0, 1
UhajSecaDruzina	Ali se je v družini pojavljalo uhajanje seča?	0, 1
PogostNehotUhajUr	Kako pogosto vam nehoteno uhaja urin?	0=nikoli, 1=1x tedensko ali manj, 2=2-3x na teden, 3= dnevno, 4=večkrat dnevno 5=ves čas
KolikoUide	Koliko urina vam uide?	0=nič, 1=majhna količina, 2=zmerna, 3=velika
NehotUhajUrMoti	Kako močno nehot. uhajanje urina moti vaš vsakdanjik?	0 - 10
ICIQ	Kratki vprašalnik za posvet o inkontinenci	0-21
Inkontinenca	Tip inkontinence	0=nima, 1=stresna, 2=urgentna, 3=mešana
ZePoiskPomoc	Ste za svoje težave kdaj poiskali strokovno pomoč?	0, 1
PogPraznSM	Pogosto praznite sečni mehur? (> 7x dnevno oz. >2x ponoči)	0, 1
PoznaVaje	Ocena poznavanja vaj (MISSINGI)	0=neustrezno, 1=delno, 2=ustrezno

NALOGA 1:

Odprite datoteko **inkontinenca.xls** in si oglejte podatke.

Odprite SPSS:

- v meniju izberite *File -> Open -> Data*
- za *Format* izberite **.xls**
- poiščite datoteko **inkontinenca.xls** in jo odprite

Opomba: V SPSSu lahko vrsto spremenljivke nastavimo v stolpcu *Measure* zavihka *Variable view*. Ta loči imenske (nominal), urejenostne (ordinal) in številske (scale) spremenljivke med sabo. Nastavitev odloča o tem, katere metode sploh lahko za posamezno spremenljivko uporabimo.

Opisna statistika

Vrste spremenljivk

- Številske spremenljivke
 - razmernostne – kjer obstaja naravna ničla (plača, teža)
 - intervalne (temperature v Celzijih, koledarska leta ipd.)
- Opisne spremenljivke
 - imenske (spol, barva las)
 - urejenostne – jih lahko uredimo v zaporedje (stadij bolezni, šolska izobrazba)

Mere centralne tendence

- Povprečje (\bar{x})
 - Vsota vrednosti, deljena z njihovim številom
 - Za številske spremenljivke
 - Vprašljiva informativnost pri asimetričnih spremenljivkah
 - Kako ga izračunamo _____
- Mediana (srednja vrednost) (med)
 - Polovica vrednosti je manjša, polovica večja
 - Neobčutljiva za izstopajoče vrednosti
 - Kako jo izračunamo _____
- Modus (najpogostejša vrednost)
 - Primeren predvsem za opisne spremenljivke
 - Nesmiseln pri zveznih številskih spremenljivkah
 - Kako ga izračunamo _____

Mere razpršenosti

- Varianca in standardni odklon (var in SD)
 - Primerni meri skupaj s povprečjem
 - Povprečni kvadrirani odmik od povprečja (variance) in koren te vrednosti (standardni odklon)
 - Kako izračunamo variance _____
 - Kako izračunamo standardni odklon _____
- Interkvartilni razmik (IQR)
 - Primeren skupaj z mediano
 - Razdalja med 25. In 75. percentilom (oziroma 1. In 3. kvartilom)
 - Kako ga izračunamo _____
- Razpon
 - Razdalja med največjo in najmanjšo vrednostjo
 - Kako ga izračunamo _____

NALOGA 2:

S pomočjo SPSSa izračunajte mere centralne tendence in mere razpršenosti za podatke o inkontinenci in jih zapišite v spodnjo tabelo.

Postopek:

Analyze -> Descriptive statistics -> Frequencies, gumb Statistics...

Opombe:

V *Frequencies* je odključana možnost prikaza frekvenčne table. Kadar možnost ni zaželenja (npr. za številske spremenljivke – ZAKAJ? - pojasnite), jo izklopite in se tako izognite predolgemu in nepreglednemu izpisu.

SPREMENLJIVKA	vrsta	\bar{x}	var	SD	med	IQR	modus	razpon
Starost								
PogostoZaprtje								
Triada								
KolikoUide								
GinekoPoseg								
ITM								

Grafični prikazi porazdelitve spremenljivk

- Histogram
 - Namenjen prikazu zveznih številskih spremenljivk (v SPSSu lahko dodate še krivuljo normalne porazdelitve)

- Na y-osi je lahko število ali delež
- Stolpični diagram (angl. barplot)
 - Namenjen prikazu diskretnih številskih ali opisnih spremenljivk
 - Na y-osi je lahko število ali delež
- Okvir z ročaji / škatla z brki (angl. boxplot, box and whiskers plot)
 - Črta na sredini je mediana
 - Spodnji in zgornji rob okvirja ustrežata 1. In 4. kvartilu
- Kolač (angl. pie)
 - Za prikaz opisnih spremenljivk z malo vrednostmi

NALOGA 3:

Izberite in z SPSSom grafično prikažite spremenljivko v tabeli zgoraj s primernim grafičnim prikazom. Izberite take spremenljivke, da boste lahko uporabili vse 4 možne grafične prikaze.

Postopek:

Graphs -> Legacy dialogs (izberite pravi grafikon)

GRAFIKON	SKICA (glede na izbrano spremenljivko)	IME SPREMENLJIVKE
histogram		
stolpični diagram		
Okvir z ročaji		
kolač		

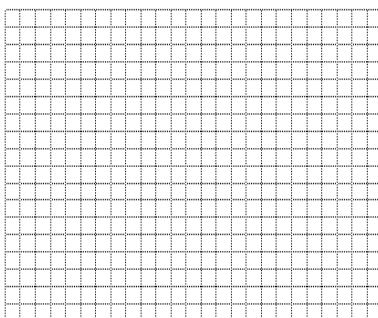
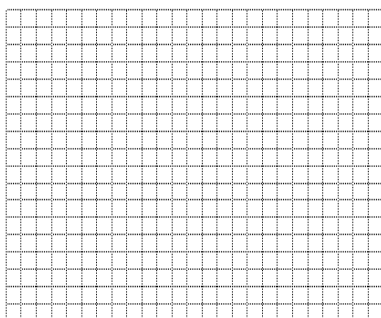
Grafični prikazi spremenljivke glede na skupine, ki jih določa druga spremenljivka

NALOGA 4:

- Stolpični diagram nehotenega uhajanja urina (PogostNehotUhajUr) glede na pogostost zaprtja (PogostZaprtje)

Graphs -> Legacy dialogs -> Bar...

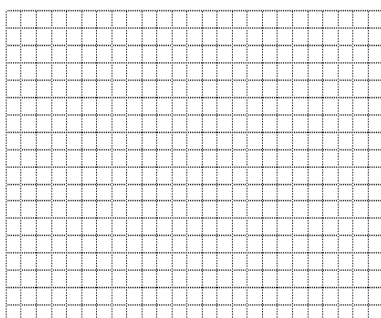
Izberemo *Summaries for groups of cases* in pogostost zaprtja kot *Rows* ali *Columns*



- Okvir z ročaji za starost glede na težave z mehurjem (PogostaVnetjaSM)

Graphs -> Legacy dialogs -> Boxplot,

Izberemo *Summaries for groups of cases* in PogostaVnetjaSM kot *Category Axis*



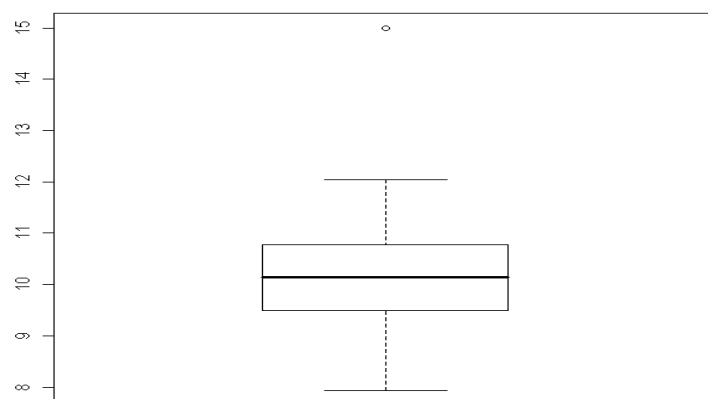
Interpretirajte oba grafikona.

Domače naloge

1. Za spodnje podatke izračunajte **modus**, **povprečje**, **mediano**, **standardni odklon** in **varianco**. Podajte razlage izračunanih statistik. Predpostavljajte, da so podatki vzorec iz neke populacije. Kaj bi se spremenilo, če bi predpostavili, da gre za podatke o populaciji (to vprašanje je sicer malo verjetno, ker ima populacija zelo redko le 8 enot!)?

Podatki (vzorec) 4,6,6,9,1,3,7,10

2. Iz spodnjega grafikona (narisali smo _____) odčitajte oziroma izračunajte: **mediano**, **1. kvartil**, **IQR**, **najmanjšo in največjo vrednost** ter **razpon**.



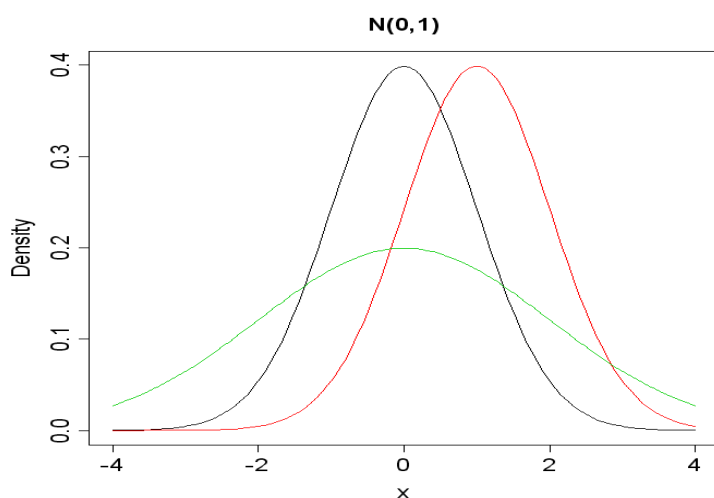
Intervali zaupanja

Normalna porazdelitev

NALOGA 5:

Zapišite vrednosti parametrov za vse tri oblike normalne porazdelitve na spodnjem grafu:

- Črna: _____
- Rdeča: _____
- Zelena: _____



CLI (centralno limitni izrek)

Povprečje neodvisnih spremenljivk, ki imajo enako porazdelitev, se porazdeli z normalno porazdelitvijo (če imamo veliko spremenljivk – velik n).

$$X_i: E(X_i) = \mu, SD(X_i) = \sigma \longrightarrow \bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Standardni odklon (SD – Standard Deviation) povprečja imenujemo **standardna napaka (SE – Standard Error)**.

NALOGA 6:

Oglejte si računalniške simulacije na spletni strani Rice Virtual Lab in Statistics – Simulations/Demonstrations: http://www.ruf.rice.edu/~lane/stat_sim/index.html Confidence Intervals. Ogleдали si bomo interval zaupanja za oceno aritmetične sredine (vpliv stopnje zaupanja in velikosti vzorca)

Če ponovite 100 krat raziskavo in vsakič izračunate 95% interval zaupanja, v koliko primerih pričakujete, da bo

- prava vrednost populacijskega povprečja zajeta v intervalih? _____
- vzorčno povprečje zajeto v intervalih? _____

NALOGA 7:

Odprite (*File* → *Open* → *Data...* ali s klikom na ustrezno ikono v opravilni vrstici) datoteko **inkontinenca2.sav**. Kakšna je standardna napaka (SE) pri vzorcih te velikosti? (Uporabite populacijski standardni odklon)

$$\sigma = 2.5 \text{ m}^2/\text{kg}, n = \text{_____} \text{ (število statističnih enot)}$$

$$SE = \sigma/\sqrt{n} = \text{_____}$$

- Ocenite 50%, 95% in 99% intervale zaupanja za oceno aritmetične sredine ITM!

Formula za izračun intervala zaupanja _____

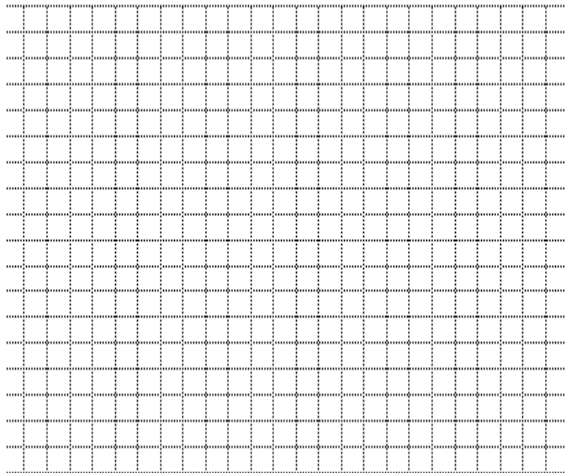
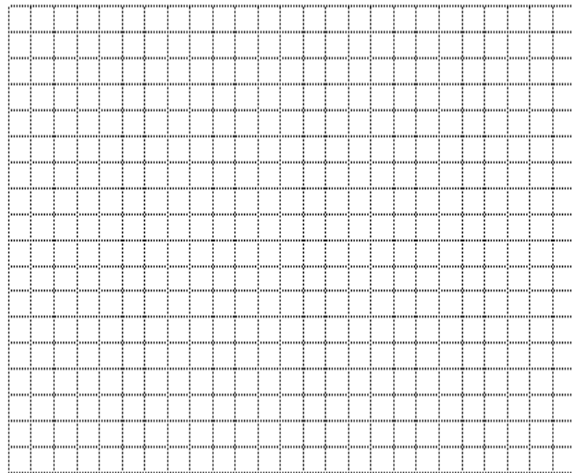
$$\bar{x} = \text{_____}$$

zaupanje	spodnja meja IZ	zgornja meja IZ
50%		
95%		
99%		

- Kateri od intervalov zaupanja je najširši? _____
- Ali ste v vse intervale zajeli pravo (populacijsko) vrednost povprečja $\mu = 24.0$? _____
- Kako se porazdeli standardizirano vzorčno povprečje, če populacijska varianca ni znana?

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim$$

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim$$

t-porazdelitev (4 d.f) in N(0,1)**t-porazdelitev (98 d.f) in N(0,1)****NALOGA 8:**

Populacijskega standardnega odklona navadno ne poznamo, zato izračune ponovite tako, da standardno napako ocenite na vzorcu:

- a) Izračunajte oceno standardne napake na podlagi vzorca.

$s =$ _____ $SE = s/\sqrt{n} =$ _____

Formula za izračun intervala zaupanja _____

- b) Ocenite meje 50, 95% in 99% intervala zaupanja za oceno aritmetične sredine ITM!

zaupanje	spodnja meja IZ	zgornja meja IZ
50%		
95%		
99%		

- c)
d) Primerjajte rezultate prve in druge tabele z IZ ter komentirajte.

Denimo, da raziskavo o ITM ponovite 100 krat in vsakič izračunate 95% IZ. V koliko primerih pričakujete, da bo

- a) populacijsko povprečje zajeto v intervalih? _____
b) vzorčno povprečje zajeto v intervalih? _____

Opomba: v SPSS je izračun IZ enostaven [postopek *Analyze* → *Descriptive Statistics* → *Explore...* v gumbu *Statistics* lahko določite stopnjo zaupanja]

DODATNO:

Kako se spremenijo IZ, če za vzorec vzamemo podatke 200 osnovnošolskih učiteljic (**inkont_200.sav**)?

zaupanje	spodnja meja IZ	zgornja meja IZ
50%		
95%		
99%		

Komentirajte rezultate.