

Obrada rezultata merenja

Milan Bjelica



Decembar 2012.

Dakle, . . .

- ▶ Naučili smo kako se koriste instrumenti,
- ▶ Izvršili smo merenja,
- ▶ Ostaje još da obradimo rezultate i izradimo izveštaj.

Čemu služi izveštaj?

- ▶ Opisuje metodologiju merenja
- ▶ Pregledno daje rezultate merenja
- ▶ Na osnovu njih daje **zaključak**
- ▶ **Ima pravnu snagu**

Šta radimo s rezultatima?

- ▶ Prikažemo ih tabelarno i grafički
- ▶ Podvrgnemo statističkoj obradi
- ▶ Uporedimo saglasnost s normama

Statistička obrada

skup rezultata = **uzorak**

obim uzorka = broj merenja

Na uzorku nalazimo min i max vrednost - očigledno kako;

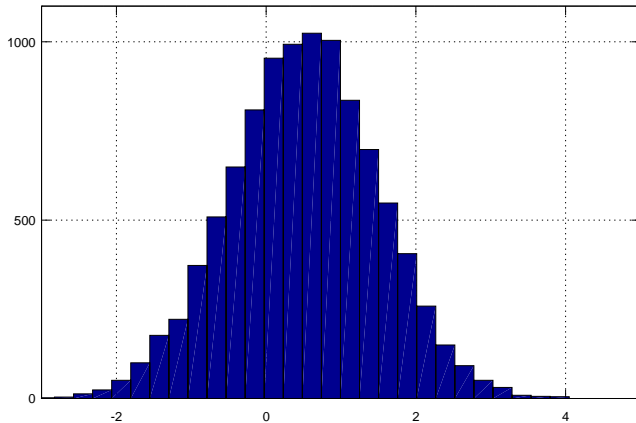
ocenjujemo matematičko očekivanje, varijansu, standardnu devijaciju, kvantile, koeficijent korelacije. . .

izvršavamo statističke testove

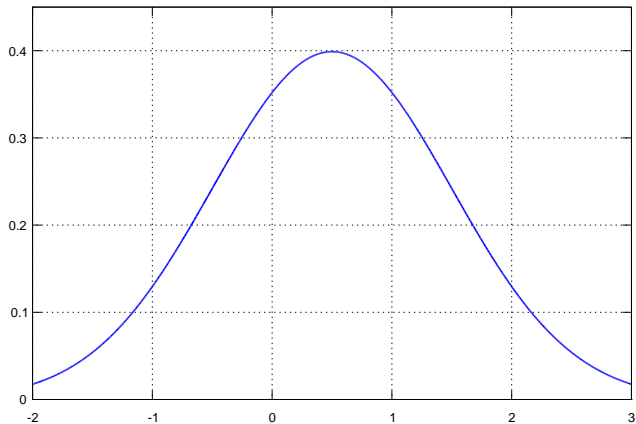
Mini podsetnik: V & S

- ▶ statistički eksperiment
- ▶ slučajna promenljiva
- ▶ histogram
- ▶ funkcija gustine verovatnoće
- ▶ funkcija raspodele

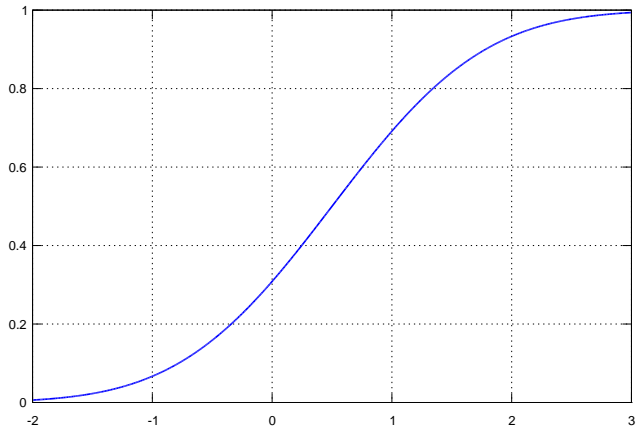
Histogram



PDF



CDF



OK, nazad na merenja

Ocena matematičkog očekivanja:

$$\hat{\mu} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Ocena varijanse

kad matematičko očekivanje *nije poznato*:

$$s^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \hat{\mu})^2 = \frac{1}{n-1} \sum_{k=1}^n X_k^2 - \frac{n}{n-1} \hat{\mu}^2$$

kad je *poznato* matematičko očekivanje:

$$s_0^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \mu)^2$$

Intervali poverenja

Def: Dvostrani IP za parameter θ s nivoom poverenja $1 - \alpha$ je interval $[Y_1, Y_2]$, za koji važi

$$P(\theta \in [Y_1, Y_2]) = 1 - \alpha.$$

Konkretno,...

$$\mathbf{P} \left(\mu \in \left[\hat{\mu} - \epsilon_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \hat{\mu} + \epsilon_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \right) = 1 - \alpha$$

Pretip:

- (1) poznata varijansa,
- (2) rezultati merenja čine nezavisan uzorak iz normalne raspodele

Za $n \geq 30$, dobra aproksimacija i za raspodele koje nisu normalne

Šta kad σ^2 nije poznato?

- ▶ umesto tačne vrednosti σ , treba staviti ocenu s
- ▶ ϵ_u je kvantil reda u iz Studentove raspodele sa $n - 1$ stepenom slobode, $t(n - 1)$ (za $n \geq 30$, svodi se na normalnu)

Dvostrani IP za variјansu

$$P \left(\sigma^2 \in \left[\frac{(n-1)s^2}{\epsilon_{1-\alpha/2}}, \frac{(n-1)s^2}{\epsilon_{\alpha/2}} \right] \right) = 1 - \alpha,$$

Pretp: rezultati merenja čine nezavisan uzorak iz normalne raspodele

ϵ_u – kvantil reda u raspodele $\chi^2(n-1)$

Šta dalje?

Završili smo deskriptivnu statistiku, sad ćemo
testirati hipoteze

Radimo testove???

Dokazivanje u statistici *nije egzaktno*, već se izvodi **testiranjem hipoteza**.

Suprotno tvrđenje (ili neutralno ili postojeće stanje) uzimamo za *nultu hipotezu*, H_0 , dok samo tvrđenje koje dokazujemo uzimamo za hipotezu H_1 .

Cilj testa: **ispitati ima li dokaza protiv H_0 , u korist H_1** .

Treba definisati statistiku testa, X i skup njenih vrednosti za koje se odbacuje nulta hipoteza.

Moguća su dva ishoda...

- (a) vrednost statistike testa unutar oblasti odbacivanja \rightarrow odbacujemo H_0 , a prihvatamo H_1
- (b) vrednost statistike testa van oblasti odbacivanja \rightarrow nema dokaza protiv H_0

...i dve vrste greške:

- (1) greška prve vrste: odbaci se H_0 , iako je tačna
- (2) greška druge vrste: H_0 se ne odbaci, iako je tačna H_1

Još malo definicija

Maksimalna vrednost greške prve vrste naziva se nivoom značajnosti testa (najčešće 0,1; 0,05; 0,025 ili 0,01)

Što je manji n_z , utoliko je za isti obim uzorka i za istu statistiku testa teže odbaciti H_0 .

Smanjivanjem n_z , povećava se verovatnoća greške druge vrste.

Pearsonov (χ^2) test

- ▶ najčešće korišćen statistički test
- ▶ svodi se na poređenje empirijskog histograma s pretpostavljenim

Testiranje hipoteza o raspodeli

Cilj: utvrditi da li se empirijska funkcija raspodele uzorka, F , podudara s nekom zadatom funkcijom raspodele, F_0 .



Kako radi ovaj test?



Verovatnoća da se i -ti element uzorka nalazi u intervalu A_j :

$$p_j = P(X_i \in A_j) = F(a_j) - F(a_{j-1})$$

Kada bi važila pretpostavljena raspodela, ova verovatnoća bi bila

$$p_{j0} = F_0(a_j) - F_0(a_{j-1})$$

Hipoteze

$$H_0 : \quad p_1 = p_{10}, p_2 = p_{20}, \dots, p_r = p_{r0}$$

$$H_1 : \quad (p_1, p_2, \dots, p_r) \neq (p_{10}, p_{20}, \dots, p_{r0})$$

Statistika testa

$$\chi^2 = \sum_{j=1}^r \frac{(N_j - np_{j0})^2}{np_{j0}} = \sum \frac{(\text{stvarno} - \text{očekivano})^2}{\text{očekivano}}$$

n – obim uzorka

N_j – broj elemenata uzorka koji se nalaze u j -tom intervalu

Odlučivanje

Nulta hipoteza se odbacuje s nivoom značajnosti α ako i samo ako je

$$\chi^2 > \epsilon_{1-\alpha}$$

$\epsilon_{1-\alpha}$ – kvantil reda $1 - \alpha$ raspodele $\chi^2(r - 1)$

Testiranje nezavisnosti

Npr: Neka se u svakom od n eksperimenata istovremeno realizuju dva ishoda. Da li su oni nezavisni?

Tablica kontingencije

	B_1	B_2	\dots	B_k	ukupno
A_1	f_{11}	f_{12}	\dots	f_{1k}	a_1
A_2	f_{21}	f_{22}	\dots	f_{2k}	a_2
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
A_v	f_{v1}	f_{v2}	\dots	f_{vk}	a_v
ukupno	b_1	b_2	\dots	b_k	n

Statistika testa

$$\chi^2 = \sum_{i=1}^v \sum_{j=1}^k \frac{(nf_{ij} - a_i b_j)^2}{na_i b_j}$$

Odlučivanje

Nulta hipoteza (o nezavisnosti događaja) se odbacuje s nivoom značajnosti α ako i samo ako je statistika testa veća od kvantila reda $1 - \alpha$ raspodele $\chi^2((v - 1)(k - 1))$.

Test Kolmogorova i Smirnova

Dok hi kvadrat test poredi histograme, K-S poredi funkcije raspodele.

Hipoteze

$$H_0 : F = F_0$$

$$H_1 : F \neq F_0$$

Statistika testa

$$\lambda = \sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$$

Odlučivanje

Nulta hipoteza se odbacuje s nivoom značajnosti α ako i samo ako je

$$\lambda > c$$

c – kvantil reda $1 - \alpha$ K -raspodele Kolmogorova

OK, dosta je bilo računanja

Kako predstaviti rezultate merenja?

- ▶ samo tabelarno (gde?)
- ▶ samo grafički (kako i gde?)
- ▶ na oba načina

Anscombe

x	y
10	8,04
8	6,95
13	7,58
9	8,81
11	8,33
14	9,96
6	7,24
4	4,26
12	10,84
7	4,82
5	5,68

x	y
10	9,14
8	8,14
13	8,74
9	8,77
11	9,26
14	8,10
6	6,13
4	3,10
12	9,13
7	7,26
5	4,74

x	y
10	7,46
8	6,77
13	12,74
9	7,11
11	7,81
14	8,84
6	6,08
4	5,39
12	8,15
7	6,42
5	5,73

x	y
8	6,58
8	5,76
8	7,71
8	8,84
8	8,47
8	7,04
8	5,25
19	12,50
8	5,56
8	7,91
8	6,89

Skoro identične statistike

Očekivanje x : 9

Varijansa x : 11

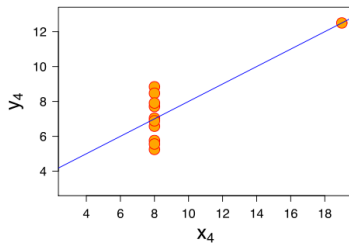
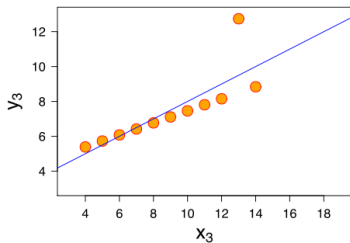
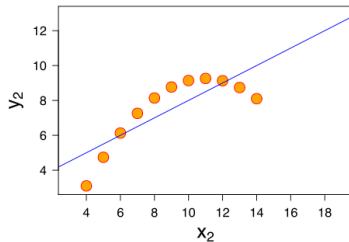
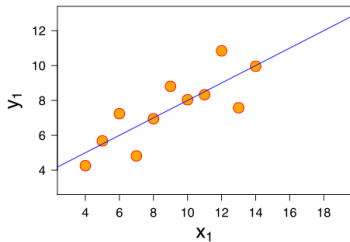
Očekivanje y : 7,50

Varijansa y : 4,122 ili 4,127

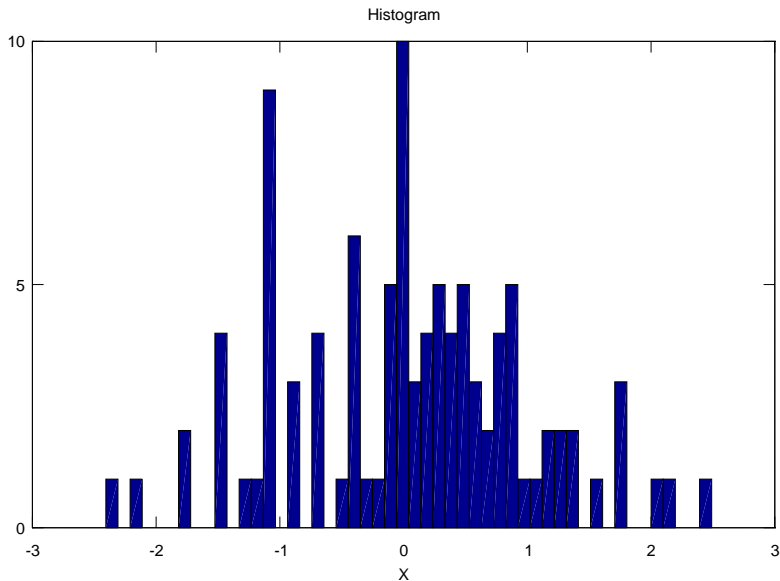
Koeficijent korelacije x i y : 0,816

Regresiona prava: $y = 3,00 + 0,500x$

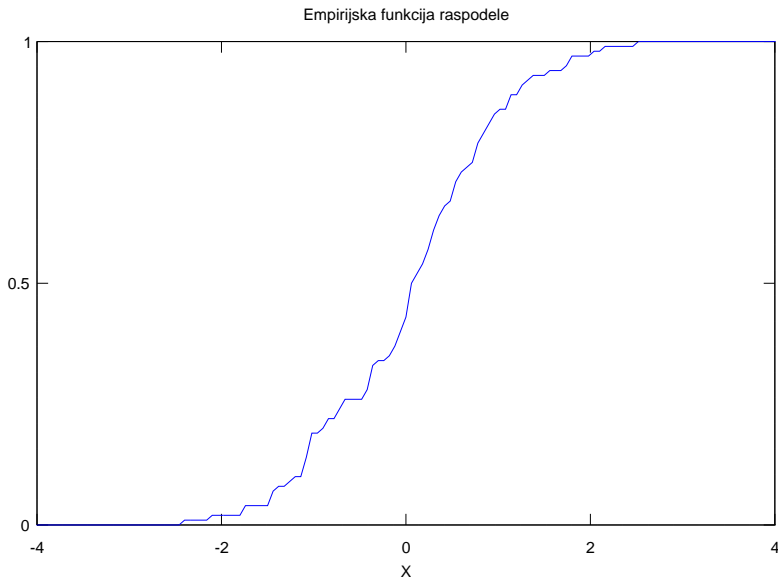
Kada nacrtamo



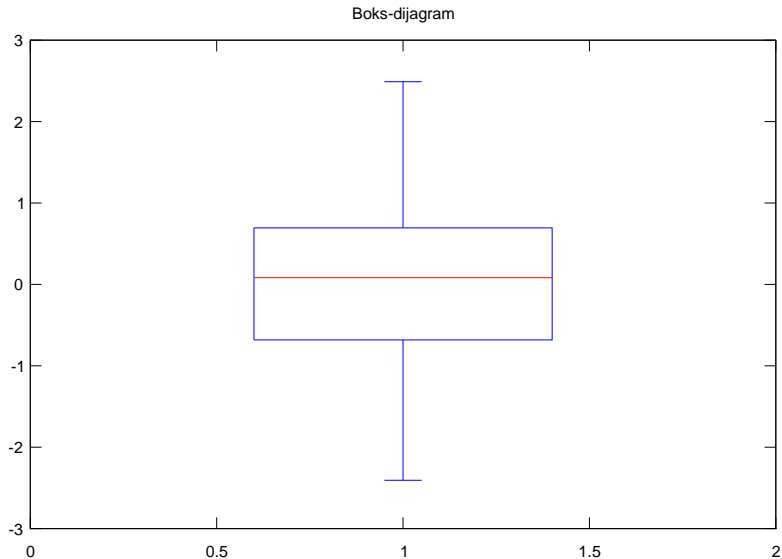
Primeri grafika: histogram



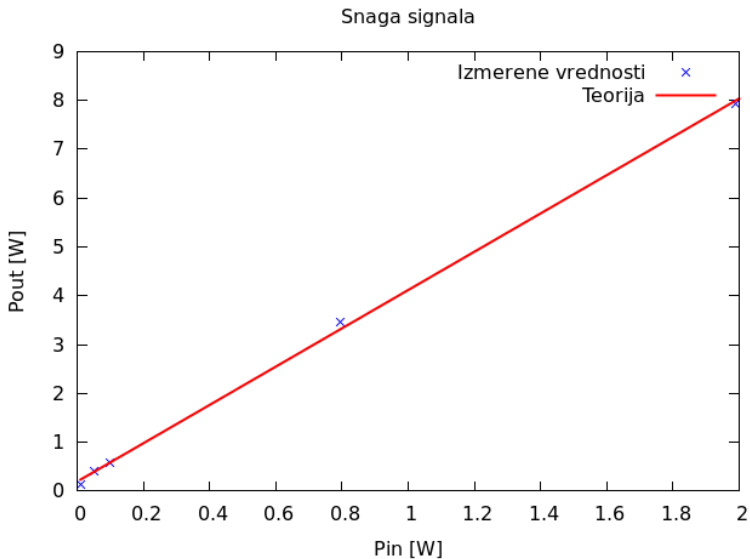
Primeri grafika: empirijska CDF



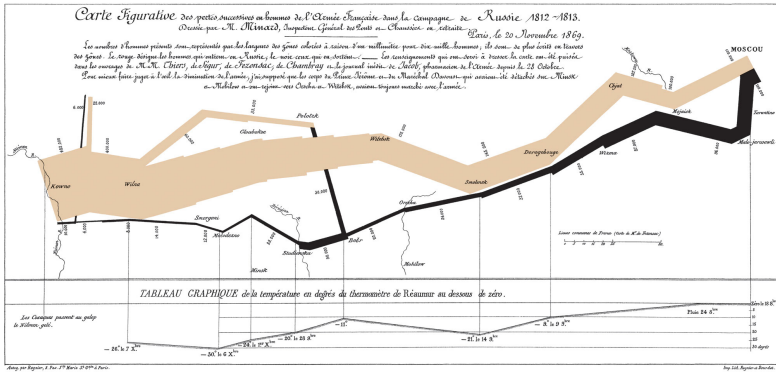
Primeri grafika: boks-dijagram



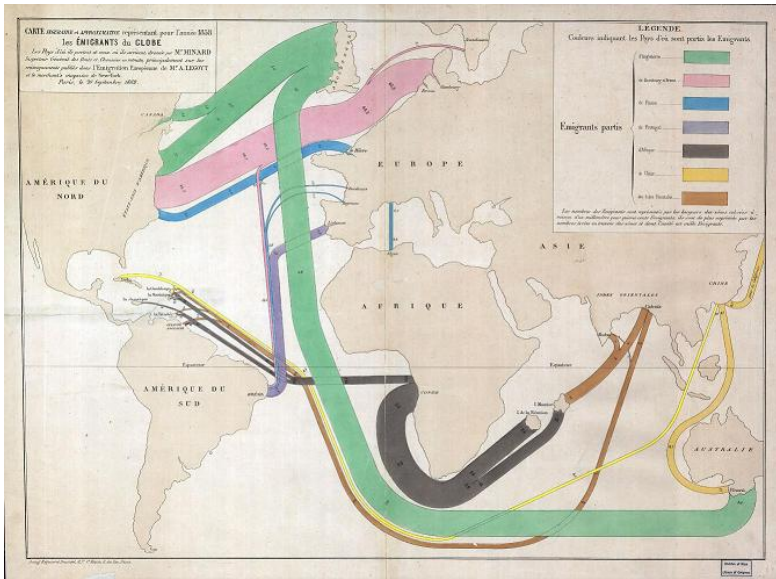
Primeri grafika: rezultati i model



Klasičan primer: Charles Joseph Minard (1)



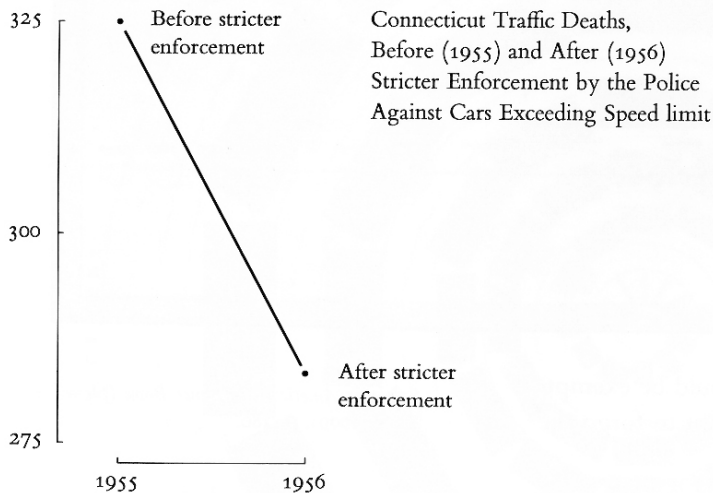
Klasičan primer: Charles Joseph Minard (2)



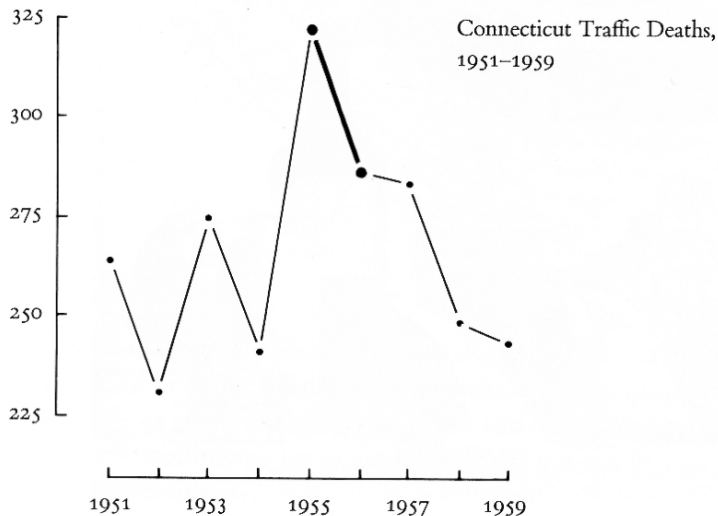
Klasičan primer: Dr. Snow



Manipulisanje graphicima (1)

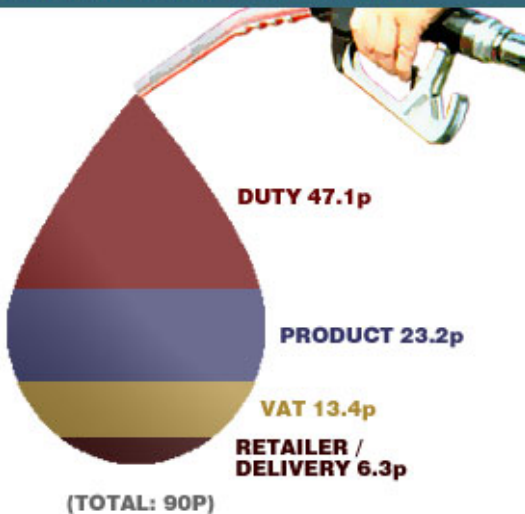


Manipulisanje graficima (2)



Manipulisanje graficima (3)

HOW ONE LITRE OF PETROL BREAKS DOWN



SOURCE: AA Sep 2005

Na kraju: propisi

Nejonizujuća zračenja: ispod 12,4 eV

Granica NF/VF: 10 kHz; 35 kV

VF: merenja na 2 godine

NF: merenja na 4 godine

Ko može da meri (pred zakonom)?

- ▶ pravno lice sa sedištem u RS
- ▶ 2 zaposlena VSS - MS, ili BS + 3 god
- ▶ akreditacija

Šta treba da sadrži izveštaj?

- ▶ opšte podatke
- ▶ opšte podatke o licima
- ▶ podatke o opremi
- ▶ podatke o merenjima
- ▶ podatke o izveštavanju