

OPTIMALNA VELIČINA UZORKA PRI ISPITIVANJU KVANTITATIVNIH SVOJSTAVA OZIME PŠENICE

Andrijana Eđed, D. Horvat, Z. Lončarić

Izvorni znanstveni članak
Original scientific paper

SAŽETAK

Prilikom planiranja pokusa važno je pravilno odrediti potrebnu veličinu uzorka, kako bi statistička analiza dobivenih podataka dala što preciznije rezultate. U radu su procijenjene optimalne veličine uzorka za ispitivanje kvantitativnih svojstava ozime pšenice (visina biljke, dužina klasa, broj klasića u klasu, broj zrna u klasu, masa zrna po klasu i masa 1000 zrna) potrebne za utvrđivanje statistički opravdanih razlika između tretmana, s pragom značajnosti od 5% i uz snagu testa od 80%.

Za izračunavanje optimalne veličine uzorka korištene su aritmetičke srednje vrijednosti i standardne devijacije kvantitativnih svojstava ozime pšenice, snaga testa, stupanj statističke značajnosti α i učinak tretmana. U radu su korišteni podaci dobiveni su iz pokusa postavljenoga po potpuno slučajnom planu, s dva tretmana u četiri ponavljanja. Na osnovi skupa od $n=440$ mjerenja, za svako ispitivano svojstvo, retrospektivno je izračunata snaga testa, nakon čega su procijenjene optimalne veličine uzorka za svako ispitivano svojstvo. Također su procijenjene optimalne veličine uzorka pri učinku tretmana od deset posto. Dobiveni rezultati pokazali su da je za svako ispitivano svojstvo različita optimalna veličina uzorka. Za utvrđivanje učinka tretmana od 10% kao statistički značajnog na razini od 5% značajnosti za broj klasića po klasu potreban je najmanji uzorak ($n = 38$), dok je najveći uzorak od $n = 440$ potreban za broj zrna po klasu i masu zrna po klasu (g) ($n = 436$). Optimalna veličina uzorka najviše ovisi o varijabilnosti ispitivanoga svojstva i veličini promjene koju želimo dokazati kao statistički opravdanu.

Ključne riječi: optimalna veličina uzorka, snaga testa, komponente prinosa, ozima pšenica.

UVOD

Određivanje optimalne veličine uzorka polazna je točka svakoga poljskoga pokusa (Confalonieri i sur., 2006.). Pokusi u području poljoprivrede najčešće se postavljaju s ciljem utvrđivanja promjene u veličini mjerenoga svojstva (ispitivane varijable) pri utjecaju različitih tretmana. S obzirom na cilj istraživanja, izabire se dizajn pokusa, broj tretmana, broj ponavljanja te se postavljaju nulta (H_0) i alternativna hipoteza (H_1). Na osnovi rezultata statističkih testova, donosi se odluka o prihvaćanju ili odbacivanju nulte hipoteze, pri čemu se mogu pojaviti dvije vrste pogrešaka. Kod odbacivanja istinite nulte hipoteze (H_0) učinjena je pogreška tipa I, koja se označava s α , a prikazuje vjerojatnost odbacivanja istinite nulte hipoteze (Tablica 1.).

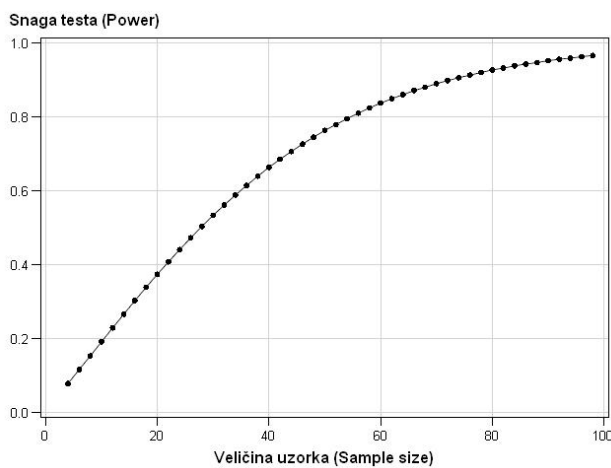
Prihvaćanjem neistinite nulte hipoteze učinjena je pogreška tipa II, koja se označava se β i predstavlja vjerojatnost prihvaćanja lažne nulte hipoteze (H_0). Uz vjerojatnost prihvaćanja lažne nulte hipoteze (β) usko je vezana snaga statističkoga testa koja se izražava kao $1-\beta$ i predstavlja vjerojatnost odbacivanja lažne nulte hipoteze. Odnos vjerojatnosti odbacivanja nulte hipoteze α i vjerojatnost prihvaćanja lažne nulte hipoteze β kao istinite temelj je od kojega se polazi pri određivanju veličine uzorka za postupak testiranja (Šošić, 2004.).

Andrijana Eđed, dipl.inž.; dr.sc. Dražen Horvat, red.prof.; dr.sc. Zdenko Lončarić, izv.prof. – Sveučilište J. J. Strossmayera u Osijeku, Poljoprivredni fakultet Osijek, Trg Sv. Trojstva 3, 31000 Osijek

Tablica 1. Tablica odlučivanja o nultoj hipotezi*Table 1. Null hypothesis decision table*

Odluka	Nulta hipoteza	
	H_0 je istinita	H_0 je neistinita
odbaciti H_0	pogreška tipa I ($\alpha = 0,05$)	ispravna odluka
prihvatiti H_0	ispravna odluka	pogreška tipa II ($\beta = 0,2$)

Veličina je uzorka u pozitivnoj korelaciji sa snagom testa (Grafikon 1.), a općenito je prihvaćeno da snaga testa ne bi trebala biti manja od 0,8 (Olbricht i Wang, 2007.) za prag značajnosti $\alpha = 0,05$. Optimalna veličina uzorka znači da uzorak treba biti dovoljno velik da zadovolji pretpostavke potrebne za izvođenje preciznih zaključaka, a opet ne prevelik zbog prevencije prihvatanja lažne nulte hipoteze (false discovery rate - FDR).

**Grafikon 1. Odnos veličine uzorka i snage testa***Figure 1. Relation of sample size and power*

Confalonieri i sur. (2006.) ističu da bi najpravičnije bilo procijeniti potrebnu veličinu uzorka na osnovi rezultata preliminarnih istraživanja, ali da se to rijetko čini. Veličine uzoraka kvantitativnih svojstava pšenice i prinosa zrna ozime pšenice u dostupnoj literaturi vrlo su različite. Maqsood i sur. (1999.) ispitali su utjecaj različitih kombinacija N-P-K gnojiva na komponente prinosa zrna i prinos zrna ozime pšenice: visina biljke (cm), duljina klasa (cm), broj klasića u klasu, broj zrna u klasu, masa 1000 zrna (g) i prinos kg ha^{-1} . Za sva ispitivana svojstva korištena je jednaka veličina uzorka od $n=24$. Gibson i Paulsen (1999.) analizirali su uzorak veličine $n=36$, Roy i sur. (2004.) uzorak veličine $n=330$, a Mian i sur. (2007.) imali su uzorak veličine $n=60$. U navedenim radovima, analizirani su podaci o vrijednostima kvantitativnih svojstava ozime pšenice i za sva ispitivana svojstva korišteni su uzorci jednake veličine. Budući da varijabilnost svih svojstava nije jednaka, potrebno je odrediti optimalnu veličinu uzorka za svako ispitivano svojstvo. Iz tog razloga, cilj ovog istraživanja bio je: (i) odrediti optimalne veličine uzorka komponenti prinosa (visina biljke, dužina klasa, broj klasića u klasu, broj zrna u klasu, masa zrna po klasu i masa 1000 zrna) na osnovi izvornih podataka – retrospektivna snaga (retrospective power), (ii) utvrditi koja je najmanja potrebna veličina uzorka svakog ispitivanog svojstva za održavanje snage testa na razini od 0,8 (80%) u provedenom pokusu, (iii) procijeniti potrebnu veličinu uzorka za utvrđivanje 10% razlike između ispitivanih tretmana za svako ispitivano svojstvo.

MATERIJAL I METODE

Podaci su prikupljeni tijekom 2008. godine iz poljskoga pokusa postavljenog u djelomično kontroliranim uvjetima po potpuno slučajnome rasporedu s četiri ponavljanja. U pokus je uključeno 55 kultivara pšenice, od kojih su 52 *Triticum aestivum*, a preostale 3 *Triticum durum*. Primjenjena su dva

tretmana (tretman 1 i tretman 2). Sjetva je obavljena u posude promjera 275 mm, u koje je odvagano 11 kg tla. Gustoća sjetve bila je 850 zrna m⁻². U svaku posudu zasijan je jedan kultivar i primjenjen jedan tretman.

Žetva je provedena u razdoblju od 23. lipnja do 1. srpnja te su zabilježeni podaci o sljedećim svojstvima: visini biljke, duljini klasa, masi biljke, masi klasa, broju klasića u klasu, broju fertilnih klasića u klasu, broju sterilnih klasića u klasu, broju zrna u klasu, masi zrna po klasu i masi 1000 zrna. U radu su analizirana svojstva: visina biljke, duljina klasa, broj klasića u klasu, broj zrna u klasu, masa zrna po klasu i masa 1000 zrna.

Statistička obrada podataka napravljena je u modulu Analyst programa SAS/STAT (SAS Institute Inc., 2002.- 2003.). Izračunate su aritmetičke sredine, standardne devijacije i koeficijenti varijacije svih svojstava (Tablica 4.) Provedena je „two sample t test analiza, na osnovi koje je utvrđeno postojanje razlika između tretmana za sva ispitivana svojstva te je određena snaga testa (retrospective power) za veličinu osnovnog uzorka n=440. Procedurom „PROC POWER“, na osnovi srednjih vrijednosti tretmana i standardne devijacije, određena je najmanja potrebna veličina uzorka svih ispitivanih svojstava za održavanje snage testa na razini od 80%.

Istom procedurom, na osnovi varijabilnosti svakoga pojedinoga ispitivanoga svojstva, određena je veličina uzorka potrebna za otkrivanje 10%-tne razlike između tretmana kao statistički značajne na razini od 0,05, uz snagu testa od 80%.

REZULTATI I RASPRAVA

Retrospektivno određivanje optimalne veličine uzorka i procjena najmanje potrebne veličine uzorka za kvantitativna svojstva ozime pšenice

Na osnovi podataka o vrijednostima kvantitativnih svojstava ozime pšenice, proveden je „two-sample t test“ (Tablica 2.). Veličina uzorka bila je n = 440 za sva ispitivana svojstva. Test je pokazao da se duljina klasa, broj klasića u klasu, broj zrna u klasu, masa zrna po klasu i masa 1000 zrna statistički visoko značajno razlikuju (P < ,0001), dok za visinu biljke nisu utvrđene statistički značajne razlike između tretmana (P = 0,886). Za isti je test napravljena retrospektivna analiza snage (retrospective power analysis) kojom je utvrđeno da za visinu biljke (cm) snaga testa iznosi 0,05, dok je za ostala ispitivana svojstva iznosila 0,99 (Tablica 2.). U literaturi se navodi da je odgovarajuća snaga testa 0,8 ili više (nikako ne bi trebala biti niža) i općenito je prihvaćeno pravilo „pet-osamdeset“, što znači da je vjerojatnost nastajanje pogreške tipa I 5%, a pogreške tipa II 20%. Prema Di Stefano (2003.), odgovarajuće vrijednosti za α i β (prema tome i za snagu testa) ne bi trebale biti fiksne, već prilagodljive svakome pojedinome pokusu i njegovim specifičnostima. Cohen (1988.) navodi da pravilo „pet-osamdeset“ treba koristiti samo onda kada se nema na osnovi čega odrediti željena snaga testa. Da bi se izračunala željena snaga testa, potrebno je odrediti veličinu uzorka, stupanj statističke značajnosti α , varijancu i učinak tretmana (Thomas, 1997.) Općenito se preporuča da se snaga testa i stupanj statističke značajnosti postave tako da vjerojatnost nastajanja pogreške s težim posljedicama bude manja.

Tablica 2. „Two-sample t test“ za kvantitativna svojstva ozime pšenice i pripadajuća snaga testa za n=440
Table 2. Two-sample t test of quantitative traits of winter wheat and observed power (n=440)

	Stupnjevi slobode <i>Degrees of freedom (Df)</i>	t-statistika <i>t statistics</i>	Pr > t <i>Pr > t</i>	Snaga testa <i>Observed power</i>
Broj klasića po klasu	438	7,07	< ,0001	0,99
Broj zrna u klasu	438	6,68	< ,0001	0,99
Duljina klasa (cm)	438	4,74	< ,0001	0,99
Masa 1000 zrna (g)	438	4,34	< ,0001	0,99
Masa zrna po klasu (g)	438	7,55	< .0001	0,99
Visina biljke (cm)	438	0,14	0,8867	0,05

Visoke vrijednosti snage testa (0,99), retrospektivno izračunate za t-test ($\alpha = 0,05$) ukazuju nam da za utvrđivanje nastale razlike između aritmetičkih sredina ispitivanih svojstava pod utjecajem tretmana,

veličina uzorka od $n = 440$ nije opravdana, odnosno da bi se nastali učinak tretmana mogao dokazati kao statistički značajan i s manjim brojem opservacija.

Očekivani učinak tretmana važan je čimbenik u određivanju potrebne veličine uzorka. Učinak tretmana treba odrediti unaprijed te, prema njemu i standardnoj devijaciji svojstva, odrediti optimalnu veličinu uzorka. Pri tome treba imati na umu da statistička značajnost ne podrazumijeva i biološku značajnost i obrnuto.

S obzirom na postignuti učinak tretmana i standardnu devijaciju svojstva, izračunate su optimalne veličine uzorka za otkrivanje nastalog učinka kao statistički značajnoga na razini od 5% značajnosti (Tablica 3.). Najveće promjene, od čak 25,64% i 22,25% u odnosu na kontrolu, zabilježene su za masu zrna po klasu (g) i broj zrna po klasu. Ta su svojstva imala i najveće koeficijente varijacije (38,03% i 37,27%). Procjenjena veličina uzorka koja predstavlja minimalnu veličinu uzorka potrebnu da bi se zabilježena razlika između tretmana mogla dokazati kao statistički značajna na razini značajnosti od 5% i uz snagu testa od 80% uz pripadajući koeficijent varijacije iznosi $n = 68$ za masu zrna po klasu i $n = 92$ za broj zrna po klasu. Najmanji učinak tretmana zabilježen je za visinu biljke (cm) i iznosi 0,19%. Da bi postignuti učinak tretmana mogao biti dokazan kao statistički značajan, veličina uzorka trebala bi biti $n = 171924$, uz pripadajuću standardnu devijaciju od 7,40 cm. Zbog nepraktičnosti, obima posla i cijene koštanja uzorci se takve veličine ne koriste u poljskim pokusima. Također, veličina promjene od 0,19%, koja u jedinici mjerenja iznosi 0,1cm (50,26 cm i 50,16 cm), ni u biološkome smislu nije od bitnoga značaja te u ovome slučaju možemo utvrditi da tretman nije značajno utjecao na ispitivano svojstvo.

U provedenom pokusu utvrđeno je da je broj klasića po klasu najmanje varijabilno svojstvo (CV = 10,66%), sa standardnom devijacijom od 1,86 klasića po klasu i učinkom tretmana od 7,66%, te je za njega procjenjena najmanja potrebna veličina uzorka od $n = 64$. Na temelju izračunate veličine uzorka i dizajna pokusa koji je uključivao dva tretmana i četiri ponavljanja, procijenjen je broj kultivara potrebnih za ispitivanje pojedinih svojstava (Tablica 3.).

Tablica 3. Procjenjene veličine uzorka kvantitativnih svojstava ozime pšenice i broj kultivara

Table 3. Estimated sample sizes of winter wheat quantitative traits and estimated number of varieties

Ispitivano svojstvo	Učinak tretmana (%) <i>Effect size (%)</i>	Procjenjena veličina uzorka <i>Estimated sample sizes</i>	Procijenjeni broj kultivara <i>Estimated number of varieties</i>
Broj klasića po klasu	7,66	64	8
Masa zrna po klasu (g)	25,64	68	9
Broj zrna u klasu	22,25	92	12
Duljina klasa (cm)	8,12	144	18
Masa 1000 zrna (g)	5,45	180	23
Visina biljke (cm)	0,19	171924	-

Procjena optimalne veličine uzorka za utvrđivanje učinka tretmana od 10% pri ispitivanju kvantitativnih svojstava ozime pšenice

Aritmetičke sredine tretmana 1 ispitivanih svojstava umanjene su za 10% vrijednosti te su na taj način izračunate očekivane aritmetičke sredine za utjecaj tretmana od 10%. Također je izračunata standardna devijacija tretmana 1, koja je, zajedno s procijenjenim aritmetičkim sredinama, poslužila za izračunavanje optimalne veličine uzorka. Podaci o aritmetičkim sredinama, standardnoj devijaciji i koeficijentu varijacije prikazani su u Tablici 4.

Tablica 4. Aritmetičke sredine tretmana, standardna devijacija tretmana 1 i koeficijent varijacije

Table 4. Arithmetic means of treatments, standard deviation of treatment 1 and coefficient of variation (%)

	¹ \bar{x}_1	² \bar{x}_2	³ \bar{x}_3	⁴ StDev ₁	⁵ CV (%)
Broj klasića po klasu	17,52	16,18	15,77	1,86	10,66
Masa 1000 zrna (g)	42,71	40,38	38,44	5,53	12,94
Visina biljke (cm)	50,26	50,16	45,23	7,40	14,73
Duljina klasa (cm)	7,24	6,65	6,52	1,25	17,37
Broj zrna u klasu	32,94	25,61	29,65	12,28	37,2
Masa zrna po klasu (g)	1,39	1,03	1,25	0,52	38,032

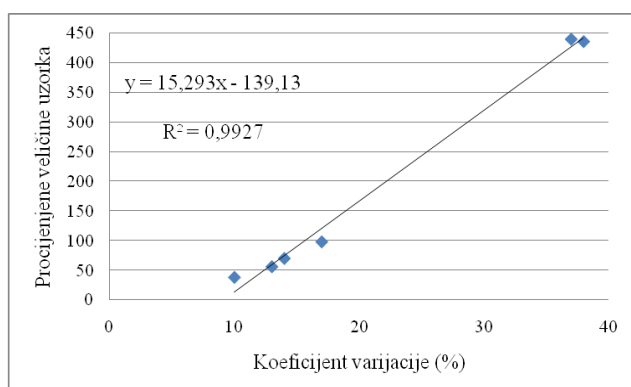
¹ aritmetička sredina tretmana 1; ² aritmetička sredina tretmana 2; ³ aritmetička sredina tretmana 1 umanjena za 10% (effect size); ⁴ standardna devijacija tretmana 1; ⁵ koeficijent varijacije (%)

Pomoću „PROC POWER“ procedure procijenjene su optimalne veličine uzoraka komponenti prinosa ozime pšenice pri učinku tretmana od 10% (Tablica 5.). Najmanji potreban uzorak, a time i najmanji potreban broj kultivara, procijenjen je za broj klasića po klasu (n = 38) i za masu 1000 zrna (g) (n = 56). Također je utvrđeno da su ta dva svojstva najmanje varijabilna s koeficijentima varijacije od 10,66% i 12,94% (Tablica 4.). Za svojstva s najvišim koeficijentima varijacije potreban je i najveći uzorak (n=436 za masu zrna po klasu (g) i n=440 za broj zrna u klasu). Između koeficijenata varijacije i veličine uzoraka za 10% promjene pod utjecajem tretmana utvrđena je visoko pozitivna korelacija (r = 0,99**) (Grafikon 2.).

Tablica 5. Procijenjene veličine uzorka kvantitativnih svojstava i broj kultivara ozime pšenice pri učinku tretmana od 10%

Table 5. Estimated sample sizes and number of varieties for winter wheat quantitative traits at 10% effect size

	Procijenjene veličine uzorka <i>Estimated sample sizes</i>	Procijenjeni broj kultivara <i>Estimated number of varieties</i>
Broj klasića po klasu	38	5
Masa 1000 zrna (g)	56	7
Visina biljke (cm)	70	9
Duljina klasa (cm)	98	13
Masa zrna po klasu (g)	436	55
Broj zrna u klasu	440	55



Grafikon 2. Scatter plot dijagram koeficijenta varijacije (%) i procijenjene veličine uzorka za učinak tretmana od 10%

Figure 2. Scatter plot diagram of coefficient of variation (%) and estimated sample sizes for 10% effect size

ZAKLJUČAK

Za izračunavanje potrebne veličine uzorka u obzir treba uzeti četiri parametra: snagu testa, stupanj statističke značajnosti α , standardnu devijaciju i učinak tretmana. Snagu testa i prag značajnosti treba odrediti prije postavljanja pokusa. Snaga testa uobičajeno se postavlja na 80% vjerojatnosti da će učinak tretmana biti uočen ukoliko stvarno postoji uz prag značajnosti od 5% vjerojatnosti da je uočeni učinak nastao slučajno. Standardnu devijaciju i učinak tretmana trebalo bi odrediti na osnovi rezultata preliminarnog istraživanja (preporuča se kao najbolja mogućnost) ili na osnovi podataka o veličinama uzoraka ispitivanih svojstava iz dostupne literature. Optimalizacija veličine uzorka, sukladno ispitivanome svojstvu, povećava vjerojatnost pouzdanosti i preciznosti dobivenih rezultata statističke analize, smanjuje obim posla kod poljskih pokusa, a time utječe na sniženje cijene koštanja ispitivanja.

ZAHVALA

Prikazani rezultati proizašli su iz znanstvenoga projekta "Utjecaj kondicioniranja tla na hraniva i teške metale u sustavu tlo-biljka" (šifra projekta: 079-0790462-0450) koji se provodi uz potporu Ministarstva znanosti, obrazovanja i športa Republike Hrvatske.

LITERATURA

1. Cohen, J. (1988.): Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates.
2. Confalonieri, R., Stroppiana, D., Boschetti, M., Gusberty, D., Bocchi, S., Acutis, M (2006): Analysis of rice sample size variability due to development stage, nitrogen fertilization, sowing technique and variety using the visual jackknife. Field Crops Research, 97:135-141.
3. Di Stefano, J. (2003): How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology, 17: 707–709.
4. Gibson, L.R. , Paulsen G.M. (1999): Yield Components of Wheat Grown under High Temperature Stress. Crop Sci., 39:1841–1846.
5. Maqsood, M., Akbar, M., Yousaf, N., Mehmood, M.T., Ahmad, S. (1999): Effect of Different Rate of N, P and K Combinations on Yield and Components of Yield of Wheat. International Journal of Agriculture & Biology, 1(4): 359 – 361.
6. Mian, M.A., Mahmood, A., Ihsan, M., Cheema, N.M. (2007): Response of different wheat genotypes to post anthesis temperature stress. J. Agric. Res., 45(4): 269–277.
7. Olbricht, G., Wang, Y. (2007.): Power and Sample Size Calculation <<http://www.stat.purdue.edu/~bacraig/SCS/>>
8. Purdue University, Collage of Science, Department of Statistics. (homepage on the Internet)
9. Roy, J. K., Lakshmikumaran, M. S., Balyan, H. S. Gupta, P. K. (2004): AFLP-Based Genetic Diversity and Its Comparison With Diversity Based on SSR, SAMPL, and Phenotypic Traits in Bread Wheat. Biochemical Genetics, 42: 43–59.
10. SAS/STAT User's Guide. (2002-2003) Version 9.1.3. Cary, NC. SAS Institute Inc.
11. Šošić, I. (2004.): Primijenjena statistika. Školska knjiga, Zagreb.
12. Thomas, L. (1997): Retrospective Power Analysis. Conservation Biology, 11(1):276-280.

OPTIMAL SAMPLE SIZE FOR STATISTICAL ANALYSIS OF WINTER WHEAT QUANTITATIVE TRAITS

SUMMARY

In the planning phase of every research particular attention should be dedicated to estimation of optimal sample size, aiming to obtain more precise and objective results of statistical analysis. The aim of this paper was to estimate optimal sample size of wheat yield components (plant height, spike length, number of spikelets per spike, number of grains per spike, weight of grains per spike and 1000 grains weight) for determination of statistically significant differences between two treatments with significance level of 5% and statistical

power of 0.8. Arithmetic means and standard deviations of winter wheat quantitative traits, statistical power, significance level α and effect size were used for calculation of optimal sample sizes.

All calculations are made on data obtained from the experiment set up according to completely randomized design with two treatments in four repetitions.

Optimal sample sizes for identification of ten percent difference between treatments are calculated for all traits. Retrospective power analysis was conducted for original data set of $n=440$ and optimal sample sizes are estimated. Obtained results showed that every examined yield component requires unique sample size. In terms of pertaining standard deviation and minimum expected difference that should be detected as statistically significant. For analyzing number of spikelets per spike sufficient sample size is 38 for effect size of 10% at significance level of 5%. Number of grains per spike and grain weight per spike (g) are the most variable traits and according to that, they have the largest estimated sample sizes ($n = 440$ and $n = 436$ respectively). Optimal sample size is mostly dependent on variability of examined trait and effect size.

Key-words: optimal sample size, statistical power, yield components, winter wheat

(Primljeno 11. svibnja 2009.; prihvaćeno 01. lipnja 2009. - Received on 11 May 2009; accepted on 1 June 2009)