

JERZY GROCHOWSKI

Rola współczynnika zmienności w pracy badawczej

The part played by the coefficient of variation in the research work

W pracy badawczej często stoimy wobec rozwiązania następujących zagadnień:

1) która z interesujących nas zmiennych cech ilościowych (np. wysokość drzew drzewostanu, liczba kształtu, iloraz q_2 itp.) odznacza się większą stałością, która zaś jest bardziej zmienną;

2) jakim średnim błędem procentowym obarczona jest wielkość określanej zmiennej, gdy badania opieramy na jednym lub więcej spostrzeżeniach,

i 3) jaka jest potrzebna liczba spostrzeżeń (wielkość próby) do określenia przeciętnej wielkości badanej zmiennej, żeby błąd, jakim będzie obarczona owa przeciętna, nie przekroczył z góry określonego średniego błędu procentowego.

Te wszystkie zagadnienia mogą być rozwiązane przy pomocy współczynnika zmienności badanej zmiennej, jeżeli rozdział jej wielkości jest normalny lub zbliżony do normalnego. Toteż wobec istotnego znaczenia wymienionych zagadnień warto współczynnikowi zmienności poświęcić nieco uwagi.

W pracach naukowych w dziedzinie leśnictwa współczynnik zmienności nie jest dość rozpowszechnionym środkiem badawczym, a nawet podręczniki statystyki matematycznej na ogół nie wiele poświęcają mu miejsca, podczas gdy ta ważna charakterystyka liczbowa jest bardzo pożytecznym narzędziem w rękach badacza ujmującego wyniki swych spostrzeżeń w liczby — w bardzo wielu dziedzinach dociekań.

W polskim piśmiennictwie dendrometrycznym spotykaliśmy się ze stosowaniem współczynnika zmienności niejednokrotnie. I tak autor niniejszej publikacji używał go do porównania zmienności wykładnika kształtu brył obrotowych, do których zbliżone są strzały drzew (śosna), ze zmiennością pierśnicy i wysokości¹⁾; B. S z y m k i e w i c z do określenia procentowego błędu, jakim może być obarczona średnia gru-

¹⁾ J. Grochowski. Wykładnik kształtu strzały. Jego rola w dendrometrii w doświadczalnictwie leśnym. Roczniki Nauk Rolniczych i Leśnych, XXIX. Poznań 1933.

bość drzew drzewostanu, w zależności od wielkości powierzchni próbnej (liczby drzew znajdujących się na powierzchni) przy znanym współczynniku zmienności grubości²⁾; T. M o l e n d a określał zmianę współczynnika zmienności pierśnicy, wysokości, liczby kształtu i liczby smukłości w zależności od wieku drzewostanów sosnowych³⁾.

Poza stosowaniem współczynnika zmienności w znaczeniu matematyczno-statystycznym, używa go się także w znaczeniu innym, a mianowicie jako miary kształtu strzały drzewa. W. P ł o ń s k i zaproponował określanie kształtu strzały za pomocą współczynnika zmienności grubości mierzonych w pewnych (krótkich) odstępach wzdłuż całej strzały⁴⁾; a autor tej pracy podał związek między wykładnikiem kształtu brył, do których zbliżone są strzały drzew leśnych, a współczynnikiem zmiennością średnic wzdłuż całej bryły⁵⁾.

Rozpatrzmy współczynnik zmienności ze stanowiska trzech wymienionych na wstępie zagadnień.

1. WSPÓŁCZYNNIK ZMIENNOŚCI JAKO MIARA WZGLĘDNEJ DYSPERSJI (ROZSIANIA⁶⁾)

Zmienność, dyspersję, rozsianie wielkości badanej zmiennej można najprościej określić za pomocą różnicy jej krańcowych wartości. Ta jednakże miara dyspersji, nie pozbawiona charakteru przypadkowego, jest tylko grubą miarą zmienności. Drugą miarą rozsiania jest przeciętne odchylenie od średniej arytmetycznej (średnia arytmetyczna bezwzględnych wartości odchyżeń poszczególnych wielkości zmiennej od wartości przeciętnej), oparte, jak wynika z definicji, na wielkościach cechy

²⁾ B. Szymkiewicz. Uwagi do pracy Prof. Dr. M. Sokołowskiego p.t. „Przyczynek do znajomości reliktywnej sosny kaukaskiej“. Sylwan. t. LV, seria A. Lwów 1937

³⁾ T. Molenda. Dendrometryczna analiza drzewostanu sosnowego w świetle badań statystycznych. Roczniki Nauk Rolniczych i Leśnych, t. XXXIX, Poznań 1937;

⁴⁾ W. Płoński. Określanie kształtu strzały drzew. Sylwan, t. LI. Lwów 1933.

⁵⁾ J. Grochowski. Współczynnik zmienności średnic jako charakterystyka kształtu strzały. Doświadczalnictwo Leśne, t. IV. Warszawa 1938.

⁶⁾ Wiadomości o współczynniku zmienności (nazywanym także wskaźnikiem zmienności), jako mierze dyspersji względnej, można znaleźć np. w następujących podręcznikach statystyki matematycznej:

1. G. U. Yule. Wstęp do teorii statystyki. Przekład Z. Limanowskiego (1921) z II wydania angielskiego (1912). Nakład Gebethnera i Wolffa.

2. J. Czekanowski. Zarys metod statystycznych w zastosowaniu do antropologii. Warszawa 1913.

3. S. Moszczeński. Podstawy organizacji gospodarstw wiejskich. Część I. Metody statystyczne w zastosowaniu do gospodarstw rolniczych, ogrodniczych i leśnych. Warszawa 1924.

4. E. Czuber. Die statistischen Forschungsmethoden. Wien 1927.

5. W. I. Romanowski. Primenienia matematicheskoj statistiki w opytnom diele. Moskwa, Leningrad 1947.

6. M. P. Ałtunin. Praktičeskoje posobie po obszčejoj i sielskochozjajstwiennoj statistike. Moskwa 1947.

7. D. Szymkiewicz. Zadania i metody statystyki. Warszawa 1948.

Także w podręczniku rachunku wyrównawczego:

8. T. M. Gołogórski. Rachunek wyrównawczy. Podręcznik dla doświadczalników i przyrodników. 1927. Nakładem Księgarni Św. Wojciecha. również w amerykańskim podręczniku dendrometrii:

9. D. Bruce and F. Schumacher. Forest Mensuration. New York and London. 1942.

(którego znaczna część poświęcona jest statystyce matematycznej)

wszystkich badanych osobników. Wreszcie najczęściej stosowaną miarą dyspersji jest odchylenie średnie: pierwiastek kwadratowy ze średniej arytmetycznej kwadratów odchyłeń od średniej arytmetycznej. Średnie odchylenie, tak samo jak wszystkie wymienione miary rozszania, jest dyspersją bezwzględną. Nie daje ono samo przez się pojęcia o stopniu zmienności danej zmiennej, bo przy takiej samej wielkości średniego odchylenia zmienność może być stosunkowo mała, jeżeli liczby szeregu statystycznego są duże i duża, jeżeli te liczby są małe. Prócz tego, przy pomocy średniego odchylenia, jako liczby mianowanej, nie można porównywać stopnia zmienności cech różnorodnych, np. liczby kształtu i wysokości drzew drzewostanu.

Do porównawczej oceny stopnia zmienności różnych cech ilościowych najodpowiedniejszą miarą jest współczynnik zmienności. Jest on średnim odchyleniem od średniej arytmetycznej wyrażonym w % % tej ostatniej. Im współczynnik zmienności jest większy, tym badana cecha jest bardziej zmienną; im mniejszy, tym odznacza się większą stałością.

W pracy pt. „Studia nad liczbą kształtu strzały drzew w drzewostanie sosnowym lasów Rogowa“, wykonanej pod kierunkiem autora niniejszej publikacji, W a c ł a w C h o j n a c k i zbadał m. in. zmienność pięciu następujących cech dendrometrycznych 536 drzew na powierzchni badawczej Zakładu Dendrometrii w lasach S. G. G. W. pod Rogowem: pierśnica, wysokość (ściślej długość strzały), pierśnicowa liczba kształtu strzały, iloraz q_2 (stosunek grubości środkowej do pierśnicy) i stosunkowa długość korony (długość korony mierzona do wierzchołka strzały od nasady pierwszej żywej gałęzi, stanowiącej ciągłość z wyżej położoną częścią korony, i wyrażona w % % długości strzały).

Powierzchnia badawcza została założona w prawie jednogatunkowym i mniej więcej jednowiekowym drzewostanie sosnowym, w wieku przeciętnym 80 lat, I bonitacji siedliska według Schwappacha, o zwarcu niejednolitym, przeważnie umiarkowanym. Wielkość powierzchni badawczej 1,5 ha. Liczba drzew 543. Na tej powierzchni poddano szczegółowym badaniom wszystkie drzewa: m. in. pomierzono pierśnice w kierunkach N—S i E—W i określono wielkości przeciętne pierśnic, pomierzono długości strzał i długości koron oraz zmierzono w środku każdej metrowej sekcji grubość w kierunkach N—S i E—W, obliczając następnie przeciętne grubości w środku sekcji. Pomiar i obliczenia wykonano: pomiar pierśnic z zaokrągleniem do 1 mm, długości strzał do 1 cm, długości koron do 1 dm, miąższość strzały obliczono sekcyjnie z zaokrągleniem do 0,0001 m³, powierzchnię przekroju pierśnicowego do 0,0001 m², liczbę kształtu z zaokrągleniem do 0,001, iloraz q_2 do 0,01 i stosunkową długość korony do 1%.

Po odrzuceniu 7 drzew (w tej liczbie 3 świerków i 2 przestojów sosnowych) W. C h o j n a c k i obliczył dla pozostałych 536 osobników współczynnik zmienności i otrzymał następujące wyniki:

q_2	h	f	$d_{1.3}$	l
6.3	6.6	7.7	18.4	20.6

(q_2 — stosunek grubości środkowej do pierśnicy, h — wysokość, f — pierśnicowa liczba kształtu strzały, $d_{1.3}$ — pierśnica, l — stosunkowa długość korony).

Z wyników tych widać, że zmienność ilorazu grubości, wysokości i liczby kształtu strzały w danym drzewostanie jest podobna i znacznie mniejsza od zmienności pierśnicy i stosunkowej długości korony. Z badań tych widać także, że liczba kształtu okazała się w tym drzewostanie bardziej zmienną niż wysokość, a największą zmiennością odznacza się stosunkowa długość korony.

Wyniki tych badań poza skonstatowaniem faktów, poza przytoczeniem się do poznania stopnia zmienności niektórych cech drzew drzewostanu sosnowego opisanego rodzaju, mają znaczenie ponadto także dla dalszych badań, o czym niżej. Mają również znaczenie praktyczne.

2. WSPÓŁCZYNNIK ZMIENNOŚCI JAKO ŚRODEK DO OCENY STOPNIA DOKŁADNOŚCI WIELKOŚCI BADANEJ ZMIENNEJ.

Jeżeli średnia arytmetyczna została obliczona z wielkości badanej cechy zmiennej u wszystkich osobników i wobec tego jest prawdziwą średnią wielkością tej zmiennej (jak to jest w przytoczonych wyżej wynikach badań współczynnika zmienności pięciu cech dendrometrycznych), to średnie odchylenie od średniej arytmetycznej jest średnim błędem pojedynczego spostrzeżenia. Zatem stosunek tego błędu do średniej arytmetycznej pomnożony przez 100, tj. współczynnik zmienności, jest średnim błędem procentowym pojedynczego spostrzeżenia.

Skoro więc współczynnik zmienności liczby kształtu strzały drzew drzewostanu wynosi 7.7, to określając liczbę kształtu na podstawie tylko jednego drzewa, wziętego z drzewostanu bez wyboru, popełniamy w stosunku do rzeczywistej przeciętnej liczby kształtu średni błąd procentowy równy 7.7%.

Prawdopodobieństwo popełnienia błędu nie większego od średniego błędu wynosi, jak wiadomo, okrągło 0.68, co oznacza, że możemy oczekiwać, iż tylko w 68 wypadkach na 100 nie popełnimy błędu większego od błędu średniego. Prawdopodobieństwo popełnienia większego błędu niż błąd średni wcale nie jest małe, bo wynosi 32 na 100. Zwykle uważa się za rzecz praktycznie pewną, że błąd pojedynczego spostrzeżenia nie przekroczy 2-3-krotnego błędu średniego, co odpowiada prawdopodobieństwu od 0.9545 do 0.9973.

Uważając, że błąd nie wykroczy poza granice 2-krotnego błędu średniego, poprzestajemy na prawdopodobieństwie 0.9545, przyjmując je za praktyczną pewność, co jednakże oznacza, że mogą się zdarzyć wypadki, iż 2-krotny średni błąd zostanie przekroczony i mianowicie okrągło w 5 wypadkach na 100. Tak samo, jeżeli utrzymujemy, że nie zostanie popełniony błąd większy od 3-krotnego średniego błędu, to nie jest to absolutna pewność, lecz tylko bardzo duże prawdopodobieństwo: okrągło 997 na 1000. Nie wykluczone jest atoli popełnienie błędu większego niż 3-krotny błąd średni, ale wypadki takie zdarzają się bardzo rzadko, bo zaledwie 3 na 1000. Przyjmowanie większego lub mniejszego współczynnika przy średnim błędzie (zwykle od 2 do 3) zależy od wagi zagadnienia i od ostrożności, z jaką do niego podchodzimy. Współczynnik większy od 3 nie jest na ogół potrzebny (prócz zagadnień wyjątkowej wagi, wymagających wyjątkowej ostrożności).

Jeżeli współczynnik zmienności liczby kształtu strzały wynosi w badanym drzewostanie 7.7%, to możemy uważać za praktycznie pewne,

że określając liczbę kształtu na podstawie jednego drzewa, wziętego bez wyboru, nie popełnimy w stosunku do prawdziwej przeciętnej liczby kształtu błędu większego od $3 \times 7.7\%$, tj. od $\pm 23.1\%$.

Ponieważ w opisywanych badaniach zwykła średnia arytmetyczna liczb kształtu wynosiła 0.418, to opierając się na jednym wziętym bez wyboru drzewie możemy otrzymać liczbę kształtu od $0.418 - 0.231 \times 0.418$ do $0.418 + 0.231 \times 0.418$, tj. od 0.321 do 0.515.

Teoretyczne te wywody znajdują potwierdzenie na drodze empirycznej. Bowiem liczba kształtu strzały drzew na powierzchni badawczej wahała się w granicach bardzo zbliżonych do powyższych, a mianowicie od 0.327 do 0.540, przy tym przedostatnia liczba kształtu w szeregu uporządkowanym według wzrastających wielkości wynosiła 0.509.

Podobnie, ponieważ współczynnik zmienności ilorazu grubości q_2 wynosi w niniejszych badaniach 6.3, to opierając się przy określaniu wartości tego ilorazu tylko na jednym drzewie, wziętym z drzewostanu bez wyboru, popełnimy w stosunku do średniej arytmetycznej ilorazów grubości wszystkich drzew średni błąd procentowy równy również 6.3%. Możemy przy tym być prawie pewni, że w stosunku do prawdziwej przeciętnej wartości ilorazów nie popełnimy błędu większego od $\pm 3 \times 6.3\%$, tj. od $\mp 18.9\%$.

Przy określaniu średniej wartości ilorazu q_2 popełnimy zatem mniejszy błąd niż przy wyznaczaniu liczby kształtu, bo zmienność (współczynnik zmienności) tego ilorazu jest mniejsza od zmienności liczby kształtu.

Jeżeli wielkość badanej zmiennej będziemy określali nie na podstawie jednego osobnika, lecz na podstawie większej liczby osobników, to średni błąd procentowy będzie oczywiście mniejszy i mianowicie, jak wiadomo, proporcjonalnie do \sqrt{n} (n — liczba osobników). Np. jeżeli przeciętną liczbę kształtu określimy na podstawie 4 drzew, to średni błąd procentowy średniej arytmetycznej (p) wyniesie:

$$p = v : \sqrt{n} = 7.7\% : \sqrt{4} = 3.85\%$$

(v — współczynnik zmienności) i możemy być prawie pewni, że otrzymana średnia liczba kształtu nie będzie się różniła od prawdziwej przeciętnej liczby kształtu więcej niż $0 \pm 3 \times 3.85\%$, tj. $0 \pm 11.55\%$.

Znając współczynnik zmienności interesującej nas cechy zmiennej możemy z góry wyznaczyć stopień dokładności, jaki osiągniemy przy określaniu średniej wartości tej zmiennej przy określonej liczbie wziętych bez wyboru (wylosowanych) osobników. Znajomość stopnia zmienności, współczynnika zmienności badanej zmiennej ma więc, jak widzimy, istotne znaczenie dla oceny dokładności, z jaką możemy wyznaczyć jej wielkość.

3. WYZNACZANIE PRZY POMOCY WSPÓŁCZYNNIKA ZMIENNOŚCI NIEZBĘDNEJ LICZBY OSOBNIKÓW DLA OSIĄGNIĘCIA Z GÓRY OKREŚLONEJ DOKŁADNOŚCI WYNIKU.

Znając współczynnik zmienności badanej zmiennej możemy z poprzedniego wzoru wyznaczyć tę liczbę osobników, przy której osiągniemy żądany stopień dokładności, z jaką ma być określona szukana wartość przeciętna. Mianowicie:

$$\sqrt{n} = \frac{v}{p}; \quad \text{skąd } n = \left(\frac{v}{p}\right)^2$$

Np. jeżeli, jak w opisanych badaniach, współczynnik zmienności liczby kształtu strzały drzew drzewostanu wynosi 7.7, to, żeby średni błąd procentowy średniej arytmetycznej liczby kształtu (p) nie przekroczył 1%, musimy wziąć do badań $n = (7.7)^2$, tj. okrągło 60 drzew, jeżeli bierzemy drzewa bez specjalnego wyboru. (Wtedy możemy być prawie pewni, że otrzymana średnia arytmetyczna będzie obciążona błędem nie większym niż $\pm 3\%$).

Do określenia przeciętnej wielkości ilorazu grubości q_2 z taką samą dokładnością (z takim samym średnim błędem procentowym) wystarczy mniej drzew, bo ta cecha jest mniej zmienna niż pierśnicowa liczba kształtu strzały. W tym wypadku liczba drzew wyniesie: $n = (6.3)^2$, tj. okrągło 40 drzew.

Jeżeli możemy zastosować odpowiedni wybór osobników reprezentujących przeciętną wielkość cechy zbioru generalnego (populacji generalnej), np. wybór drzew pod względem przeciętnej wysokości drzewostanu jednowiekowego, wówczas powiększa się w porównaniu z podanymi wzorami dokładność wyniku przy takiej samej liczbie badanych osobników lub też zmniejsza się liczba osobników przy takiej samej żądanej dokładności.

Szczególnego znaczenia nabierają wzory na dokładność przeciętnej i na liczbę osobników niezbędną do badań właśnie wtedy, gdy brak dostatecznie pewnych kryteriów przy ocenie reprezentacyjności wybranych osobników pod względem badanej cechy i gdy wobec tego wybór jest, jeżeli nie całkowicie, to w dużej mierze rzeczą przypadku. W tym wypadku wybór sprowadza się właściwie do losowania. Tak jest właśnie, w każdym razie w dużym stopniu, z wyborem drzew próbnych pod względem liczby kształtu lub pod względem przyrostu, bo wybierający drzewa próbne nie ma dostatecznych podstaw do twierdzenia, że wybór drzew jest trafny.

Wniosek: do określenia liczby osobników niezbędnych do badań i do oceny stopnia dokładności wyniku wielce pomocną jest znajomość zmienności badanej zmiennej, znajomość bodaj rzędu jej zmienności. Na tym właśnie polega główne znaczenie współczynnika zmienności dla pracy badawczej.

Dlatego też jest rzeczą bardzo celową rozszerzenie badań współczynnika zmienności na inne ważne cechy drzew, a zwłaszcza na te cechy, które są przedmiotem szczególnego zainteresowania. Z cech dendrometrycznych zaliczyć do nich należy przede wszystkim przyrost bieżący.

Wspólnym wysiłkiem Zakładu Dendrometrii S. G. G. W. i Zakładu Dendrometrii Instytutu Badawczego Leśnictwa wykonano na powierzchni badawczej, wspomnianej w niniejszej publikacji, w lasach Szkoły Głównej Gospodarstwa Wiejskiego szczegółowe pomiary także w zakresie przyrostu grubości, wysokości i miąższości strzały wszystkich drzew. Da to możliwość m. in. zbadania współczynnika zmienności również cech przyrostowych. Rzuci on światło na stopień zmienności tych cech w porównaniu z innymi cechami, a także na liczbę osobników niezbędnych do badań oraz na osiąganą dokładność w badaniach przyrostu.

SUMMARY

The author discusses the coefficient of variation as a means for determination: 1) of the degree of variation of the variable to be investigated, 2) of the degree of accuracy of the value of variable to be determined and, 3) for determination of the number of observations necessary to insure a given degree of accuracy.

In the „Studies of the form factor of the stem of the trees in a pine stand of the Rogów Forest“ made under the direction of the author of this paper by **W. Chojnacki** the following values of the coefficient of variation (v) are found out.

for	q_2	h	f	$d_{1.3}$	
v	6.3	6.6	7.7	18.4	20.6

(q_2 — the form quotient being the ratio of the diameter taken at a point one — half the height of the tree to diameter breast high, h — the height, f — stem-form factor, $d_{1.3}$ — diameter breast high, l — length of the tree crown expressed as a percentage of the tree height).

The basis of the research have been the measurements of all the 536 trees being on the 1.5 ha sample plot of Rogów Forest belonging to College of Agriculture in Warsaw. The measurements have been executed (except of the diameter breast high) after the felling of all the trees, the length of the sections have been 1 meter. The results of the investigations relate to a nearly even-aged pine stand, the average age 80 years, quality of locality I (Schwappach), the degree of crown-contact not always the same, in the most part forming a full canopy.

The numbers show that the lowest relative dispersion has the form quotient, the highest the relative length of the crown.

On the basis of the results of this research the percentage standard error of the average stem-form factor of the similar stands may be calculated depending on the number of the measurements of the trees. For example: if the stem-form factor has been determined on the basis of one tree only, its standard error amounts (nearly) to 7.7%; if on 4 trees — 3.85%, etc. The percentage standard error of the form quotient is lower by the same number of measurements, because the coefficient of variation of the form quotient is lower, than that of the form factor.

The results of this research give the possibility to calculate the number of measurements necessary to insure a given degree of accuracy. For example: if the percentage standard error should not be higher than 1%, the average stem-form factor must be determined in similar stands on the basis of about 60 trees, the average form quotient on about 40 trees.

As the knowledge of the magnitude of the coefficient of variation is important not only for science, but also for practice, it would be useful to extend the researches on its value, the more so, with reference to the increment.