

## ZMIENNOŚĆ I POWTARZALNOŚĆ OCEN CECH POKROJOWYCH U NOREK STANDARD \*

*Grażyna Jeżewska, Janusz Maciejowski*

Instytut Biologicznych Postaw Produkcji Zwierzęcej AR w Lublinie

*Jerzy Sławoń*

Ośrodek Badawczo-Rozwojowy PPL „Las” w Skolimowie

Skuteczność selekcji na cechy ilościowe, wyrażająca się postępowaniem hodowlanym, stoi w prostym stosunku do odziedziczalności cech oraz intensywności selekcji. W przypadku cech mierzalnych określenie intensywności selekcji nie następuje większych kłopotów, wyraża się ona bowiem w formie różnicy selekcyjnej. Cechy oceniane subiektywną metodą punktową, a do takich należy większość cech pokrojowych zwierząt futerkowych, nie mogą dać jednoznacznych wskazań. Tutaj na intensywność selekcji wpływa w poważnym stopniu dokładność oceny, zależna od kwalifikacji osoby oceniającej. Wcześniejsze wyniki badań [1] ustaliły, że ferowane przez poszczególnych sędziów oceny przy licencji młodych zwierząt dość znacznie odbiegają od siebie, a statystycznie istotna interakcja sędzia  $\times$  zwierzę dowiodła, że niemal każdy oceniający inaczej nieco wyobraża sobie idealne, pod względem wszystkich ocenianych cech zwierzę. Błąd oceny powiększany jest dodatkowo przez zmienność orzeczeń sędziego w stosunku do tego samego zwierzęcia.

Autorzy, stosownie do sugestii Centralnej Stacji Hodowli Zwierząt podjęli próbę określenia stopnia zmienności ferowanych ocen przy licencji nerek odmiany standard w celu znalezienia jednolitego wskaźnika liczbowego, który mógłby być użyty jako miernik kwalifikacji osób oceniających.

---

\* Praca wykonana w ramach Problemu Międzyresortowego MR-II/9 koordynowanego przez Polską Akademię Nauk.

## MATERIAŁ I METODY

Testowi sprawdzającemu poddani zostali sędziowie licencyjni o różnym stopniu doświadczenia w łącznej liczbie 14 osób. Zostali oni podzieleni na dwie grupy 7-osobowe, z których każda oceniała indywidualnie 15 nerek odmiany standard. Każde zwierzę było oceniane przez każdego sędziego trzykrotnie, przy czym zwierzęta podawano do oceny w niezna-nej sędziom kolejności. Tak więc każde zwierzę oceniono 21 razy (7 sędziów  $\times$  3 oceny), a każdy sędzia wydał 45 ocen (15 zwierząt  $\times$  3 oceny). Grupa druga oceniała inne zwierzęta, istniała bowiem obawa, że to samo zwierzę może nie przetrzymać fizycznie ponad 40 ocen. Zwierzęta oceniono zgodnie z obowiązującym wzorcem oceny pokroju [3] pod względem 5 cech, z których każda mogła uzyskać maksymalnie 6 punktów. Ponieważ pierwsza z cech — masa ciała mierzona jest zawsze obiektywnie przy pomocy wagi — nie zależy od kwalifikacji sędziego, umownie przyjęto, że cecha ta u wszystkich ocenianych zwierząt spełnia wymagania wzorca, przyznając każdej norce po 6 punktów.

Test przeprowadzono w okresie pełnej dojrzałości okrywy włosowej nerek, w typowych dla licencji warunkach fermowych, przy pogodzie charakterystycznej dla późnej jesieni — zachmurzeniu dużym, przejściowo umiarkowanym.

Wyniki oceny poddano analizie wariancji w dwukrotnej klasyfikacji krzyżowej [2], oddzielnie dla każdej z dwóch serii ocen. Ustalając komponenty wariancyjne, oszacowano w każdej serii współczynnik powtarzalności ocen. W celu porównania kwalifikacji testowanych sędziów przeprowadzono również analizę wariancji dla 45 ocen każdego sędziego, ustalając komponenty wariancji między ocenianymi zwierzętami i wewnątrz zwierząt (między kolejnymi ocenami tego samego zwierzęcia). Dało to podstawę do oszacowania współczynnika powtarzalności ocen każdego sędziego.

## WYNIKI I DYSKUSJA

W tabeli 1 podano wyniki ocen wszystkich zwierząt ocenianych w obu seriach (grupach). Podano średnie wyniki uzyskane od 7 sędziów w trzech kolejnych ocenach oraz wynik średni ze wszystkich ocen wszystkich sędziów. Tabela 1 informuje również o rozpiętości indywidualnych ocen od najniższej do najwyższej dla każdego zwierzęcia. Wyniki te są bardzo interesujące, wskazują bowiem na stosunkowo duże różnice ocen między zwierzętami i wysoką powtarzalność średnich ocen w kolejnych przeglądach zwierząt. Należy jednak pamiętać, że ta wysoka zgodność kolejnych ocen jest w dużym stopniu wyrównana przez obliczenie średnich arytmetycznych z sądów wszystkich oceniających. O stopniu zmien-

Tabela 1

Średnie oceny zwierząt w serii I i II uzyskane w kolejnych podaniach i łącznie

Lp.	Nr zwierzęcia	Płeć	Oceny				Rozpiętość ocen indywidualnych
			I	II	III	łącznie	
Seria I							
1.	1731	♂	27,71	27,57	27,86	27,71	25—29
2.	1756	♀	27,14	26,85	27,57	27,19	25—29
3.	1733	♂	27,71	27,71	27,85	27,76	24—30
4.	1174	♀	24,28	23,57	22,57	23,52	17—27
5.	1172	♀	26,85	27,28	27,85	27,33	25—29
6.	1150	♀	29,0	29,57	29,42	29,33	27—30
7.	1121	♂	27,71	27,14	27,42	27,43	24—29
8.	2255	♂	26,14	26,42	25,85	26,14	24—29
9.	2294	♀	24,28	23,0	24,14	23,81	18—28
10.	2161	♂	24,0	23,71	22,57	23,43	19—27
11.	2208	♀	22,0	20,57	21,0	21,19	16—26
12.	2165	♂	26,71	25,28	25,00	25,67	20—29
13.	2253	♂	19,71	21,42	20,71	20,62	16—27
14.	2298	♀	22,57	22,14	22,71	22,48	18—26
15.	2296	♀	25,57	26,71	26,71	26,33	23—28
Ogółem			25,42	25,28	25,28	25,33	
Seria II							
1.	1700	♀	29,0	28,57	28,85	28,81	28—30
2.	1702	♀	27,42	28,71	28,0	28,05	24—30
3.	1665	♂	26,28	26,28	25,85	26,14	22—28
4.	1822	♀	24,71	25,85	26,14	25,57	21—29
5.	1820	♀	29,0	26,0	25,42	26,81	24—29
6.	1824	♀	26,85	26,0	26,71	26,52	25—28
7.	1046	♀	28,0	27,57	27,42	27,67	25—30
8.	1015	♂	26,14	28,57	28,57	27,76	25—30
9.	1044	♀	23,85	23,14	23,14	23,38	20—29
10.	634	♂	19,42	17,28	16,85	17,85	6—24
11.	638	♀	26,28	26,0	25,71	26,00	23—28
12.	636	♀	24,28	24,42	23,57	24,09	21—28
13.	954	♀	24,85	24,42	23,57	24,28	21—29
14.	945	♂	23,85	24,42	24,28	24,19	21—28
15.	832	♀	24,85	25,00	24,57	24,81	22—26
Ogółem			25,65	25,48	25,24	25,46	

ności ocen indywidualnych dają pojęcie kolumny „rozpiętości ocen”. Tutaj różnice między oceną maksymalną i minimalną są bardzo wysokie, zwłaszcza w stosunku do niektórych zwierząt. Z przedstawionych wyników należy sądzić, że obowiązujące przepisy oceny pokroju pozwalają w sposób zadowalający wyróżnić zwierzęta lepsze od gorszych, trafność

## Analiza wariacji i powtarzalność ocen w obu seriach doświadczenia

Seria	Źródło zmienności	Stopnie swobody	Suma kwadratów odchyłeń	Średni kwadrat	Komponenty wariacji	Udział komponentu w całości wariacji %	Powtarzalność ocen
I	Między sędziami (A)	6	229,84	38,307**	$\sigma_e^2 + (15 \cdot 3) \sigma_A^2$	6,65	$\frac{\sigma_A^2 + \sigma_B^2}{\sigma_A^2 + \sigma_B^2 + \sigma_{AB}^2 + \sigma_e^2} = 0,642$
	Między zwierzętami (B)	14	2056,89	146,921**	$\sigma_e^2 + (7 \cdot 3) \sigma_B^2$	57,60	
	Interakcja A × B	84	632,92	7,535**	$\sigma_e^2 + 3 \sigma_{AB}^2$	13,71	
	Reszta	210	552,02	2,629	$\sigma_e^2$	22,04	
	Całość	314	3471,67	—	—	100,0	
II	Między sędziami (A)	6	230,34	38,390**	$\sigma_e^2 + (15 \cdot 3) \sigma_A^2$	6,65	$\frac{\sigma_A^2 + \sigma_B^2}{\sigma_A^2 + \sigma_B^2 + \sigma_{AB}^2 + \sigma_e^2} = 0,648$
	Między zwierzętami (B)	14	2083,38	148,813**	$\sigma_e^2 + (7 \cdot 3) \sigma_B^2$	58,18	
	Interakcja A × B	84	629,22	7,491**	$\sigma_e^2 + 3 \sigma_{AB}^2$	13,72	
	Reszta	210	539,40	2,568	$\sigma_e^2$	21,45	
	Całość	314	3482,34	—	—	100,0	

\*\* Różnice istotne z ryzykiem błędu 1%.

jednak oceny zależy w każdym przypadku od kwalifikacji oceniającego. Wniosek ten potwierdza przeprowadzona analiza wariancji (tab. 2), która dla obu serii doświadczenia wykazała wysokoistotne różnice między ocenianymi zwierzętami, ale również różnice między sędziami, a także potwierdziła istotność interakcji sędzia  $\times$  zwierzę. Oszacowane komponenty wariancyjne pozwoliły również na ustalenie wskaźników powtarzalności dla obu serii doświadczenia. Ich wielkość jest niemal identyczna.

Ponieważ łączny współczynnik powtarzalności daje jedynie ogólny pogląd na wiarygodność oceny, postanowiono dodatkowo oszacować oddzielnie dla każdego sędziego wskaźnik powtarzalności ocen, opierając się na analizie wariancji 45 wydanych przez każdego z nich ocen. Stosunek komponentu wariancji między zwierzętami do ogólnej wariancji był szacunkiem powtarzalności. Wyniki zamieszczono w tabeli 3. Obok powtarzalności zamieszczono w tabeli różnicę między średnią oceną każdego sędziego a średnią ogólną.

Zmienność wysokości wskaźnika powtarzalności ocen jest dosyć wysoka i zawiera się w granicach od  $r=0,399$  do  $r=0,899$ . Bliższa analiza arkusza ocen każdego sędziego i uzyskanego współczynnika powtarzalności dostarcza interesujących spostrzeżeń, które wskazują, że obliczony wskaźnik jest dobrą i obiektywną miarą kwalifikacji sędziego. Jego wielkość zależy nie tylko od zgodności ocen tego samego zwierzęcia, ale i od różnic ferowanych ocen w stosunku do różnych zwierząt. Tak np. sędzia 2 dość trafnie powtarzał oceny w kolejnych oględzinach zwierzęcia, ale na ogół większość zwierząt punktował dość wysoko i w wąskich granicach, w związku z czym współczynnik powtarzalności jego ocen osiągnął

Tabela 3

## Indywidualne średnie wyniki ocen i ich powtarzalność

Seria I				Seria II			
nr sędziego	średnia ocena pokroju	odchylenie od średniej ogólnej	powtarzalność ocen	nr sędziego	średnia ocena pokroju	odchylenie od średniej ogólnej	powtarzalność ocen
1	25,78	+0,45	0,522	8	24,44	-1,02	0,833
2	26,64	+1,31	0,589	9	24,07	-1,39	0,899
3	24,40	-0,93	0,890	10	26,09	+0,63	0,612
4	24,29	-1,04	0,898	11	25,82	+0,36	0,680
5	26,27	+0,94	0,763	12	26,71	+1,25	0,769
6	24,75	-0,58	0,692	13	25,42	-0,04	0,399
7	25,18	-0,15	0,655	14	25,69	+0,23	0,783
Ocena ogólna dla serii	25,33	—	0,643		25,46	—	0,648

wartość niższą od średniej ogólnej. Można stwierdzić, że jego kwalifikacje nie są zbyt wysokie, nie różnicuje on bowiem dostatecznie wyraźnie zwierząt lepszych od gorszych, a zatem ferowane przez niego oceny mają mierną wartość selekcyjną. Przeciwnieństwem dla sędziego 2 są wyniki uzyskane przez sędziów nr 3, 4 i 9. Sięgające prawie  $r=0,9$  współczynniki powtarzalności powstały jako wynik dużego zróżnicowania między ocenianymi zwierzętami i dużej powtarzalności tych ocen w kolejnych oględzinach tych samych zwierząt. Kiedy jednak sędzia nr 13 zróżnicował również bardzo znacznie swoje oceny, punktując jedne zwierzęta bardzo wysoko a inne bardzo nisko, a w kolejnych ocenach nie potrafił powtórzyć tych ocen, uzyskał najniższy wskaźnik powtarzalności  $r=0,399$ .

Wydaje się, że powtarzalność ocen oparta na trzykrotnym przegłądzie 15 zwierząt jest zupełnie wystarczającym miernikiem kwalifikacji sędziego licencyjnego i nadawanie uprawnień sędziowskich powinno być poprzedzone egzaminem praktycznym, którego ocenę powinien stanowić współczynnik powtarzalności. Jako dolną granicę, przekroczenie której dawałoby uprawnienia sędziego licencyjnego, należałoby przyjąć uzyskaną przeciętną  $r=0,65$ , chociaż w zasadzie przydatne dla selekcji wyniki ocen wymagałyby może podniesienia tej bariery do  $r=0,7$ .

W tabeli 3 zawarte są dodatkowe informacje, które mogą rzutować na ocenę kwalifikacji sędziego. Są to różnice wyrażające odchylenie średniej oceny pojedynczego sędziego od średniej ogólnej. Sięgają one u niektórych sędziów, którzy uzyskali wysoką powtarzalność ocen, powyżej jednego punktu, co wskazuje na duży brak zgodności oceny między sędziami. Wydaje się, że problem ten w tabeli 3 jest nieco przejawskrawiony, bowiem średnia ocena została tu obliczona na podstawie wszystkich ocen testowanych osób. Gdyby uznać osoby, które nie uzyskały  $r=0,7$  jako nie mające dostatecznych kwalifikacji, a średnią arytmetyczną ocen obliczyć tylko z wyników sędziów o wysokiej powtarzalności, różnice — jak można sądzić z tabeli — znacznie by się zmniejszyły.

#### PODSUMOWANIE I WNIOSKI

Przeprowadzony test na zgodność ocen w kwalifikacji hodowlanej nerek standard ujawnił przy pomocy rachunku statystycznego, że opierając się na obowiązujących kryteriach oceny zwierząt futerkowych można wyróżnić zwierzęta lepsze od gorszych. Stwierdzono również statystycznie istotne różnice między sędziami, co wskazuje na zróżnicowanie ich kwalifikacji, a także statystyczną istotność interakcji sędzia  $\times$  zwierzę.

Obliczone wskaźniki powtarzalności ocen z łącznej analizy wariancji i z indywidualnych ocen każdego testowanego wskazują, że są one obiek-

tywnym miernikiem umiejętności sędziego i mogą być z powodzeniem stosowane jako kryterium kwalifikacyjne przy nadawaniu uprawnień sędziowskich.

#### LITERATURA

1. Maciejowski J., Sławoń J.: Subiektywność licencyjnych ocen norek standard jako źródło błędu w pracy hodowlanej. Pr. Mater. Zoot., 4, 1973, 69-81.
2. Ruszczyc Z.: Metodyka doświadczalnictwa zootechnicznego. PWRiL Warszawa 1970.
3. Wzorce oceny pokroju lisów pospolitych, lisów polarnych i norek. Min. Roln., PWRiL Warszawa 1962.

*Г. Ежевска, Я. Мацеёвски, Е. Славонь*

#### ИЗМЕНЧИВОСТЬ И ПОВТОРЯЕМОСТЬ ОЦЕНОК ПРИЗНАКОВ ГАБИТУСА У СТАНДАРТНЫХ НОРОК

##### Резюме

Опубликованные результаты прежних исследований авторов показали, что даваемые отдельные оценки при лицензии молодых пушных зверей значительно отличаются друг от друга. Статистически достоверное взаимодействие: судья  $\times$  животное свидетельствует о том, что почти каждый оценивающий несколько иначе выобразает себе идеальное по отношению всех признаков животное. В результатах исследований заинтересована Центральная станция разведения животных, в связи со стремлением к унификации критериев оценки, а следовательно и селекции.

Осенью 1978 г. было проведено испытание 14 зоотехников Станции на повторяемость оценок. Испытание проводилось в двух сериях по 7 человек. В каждой серии испытуемые зоотехники трехкратно оценивали 15 одних и тех же животных, показываемых в неизвестной для них очередности. Оценка проводилась в соответствии с действительным эталоном оценки габитуса. Оценивали 4 признака (чистоту окраски, густоту и длину волоса и общий внешний вид). В связи с тем, что пятый признак — вес тела, определяется объективно путем взвешивания, было условно принято, что все животные отвечают в данном случае эталону, признавая им за этот признак максимальное число 6 баллов. Каждый признак у отдельных животных оценивался таким образом 21 раз (7 судей  $\times$  3 оценки).

На основании суммы оценок исчисляли среднюю арифметическую, принимая эту величину как приближенную к действительному качеству данного признака. Была принята предпосылка, что отклонения в оценке как на плюс, так и на минус, являются равно вероятными. Эти средние составляют в оценке судей центральный пункт, с которым сравниваются индивидуальные оценки. Если несогласованность собственных очередных оценок может свидетельствовать о малом опыте, недостаточной наблюдательности или слабом зрении судьи, то

отклонения индивидуальных оценок от средней оценки указывают на различное понимание эталона оценки, что создает необходимость постоянной учебы судей для унификации критериев оценки.

Авторы, основываясь на полученных материалах, пытаются разработать числовые указатели, которые составляли бы минимальные порог изменчивости оценок, преодоление которого обеспечивало бы права лицензионного судьи.

*G. Jeżewska, J. Maciejowski, J. Sławoń*

## VARIABILITY AND REPEATABILITY OF ESTIMATES OF EXTERIOR FEATURES OF MINKS OF THE STANDARD VARIETY

### Summary

In earlier investigations of the authors it has been found that the estimates given by licence judges to young fur animals considerable deviated one from another. A statistically significant interaction: judge  $\times$  animal proved that almost every judge had somewhat another idea of the animal with regard to all features estimated. The Central Animal Breeding Station is interested in results of the respective testing in view of the trend to unification of estimation criteria and consequently of selection.

In autumn 1978 14 zootechnicians of the Station were tested for the repeatability of estimates. The test was carried out in two series by 7 people. In every series the zootechnicians tested estimated three times 15 the same animals presented in a succession unknown to them. The estimation was carried out in accordance with the valid habitus estimation pattern. Four habitus features — colour purity. Since the fifth feature — body weight, is determined objectively by weighing, it was assumed conventionally that all the animals would correspond in this respect with the pattern and were given for this feature the maximum score 6. Thus each feature was estimated for each animal 21 times (7 judges  $\times$  3 estimates).

The arithmetic mean was calculated basing on the sum of the above estimates, while assuming this value as approximate to the actual value of the given feature. It was assumed that deviations in estimate, both in plus and in minus, would be equally probable. These means constitute a central point in the estimation of judges, with which individual estimates can be compared. While the incompatibility of subsequent own estimates can bear evidence of a lacking experience, insatisfactory perceptivity or weak sight of the given judge, so the divergence of individual estimates from mean would prove a different idea of the estimation pattern, what suggests the need of a continuous training of judges for unification of the estimation results.

The authors, basing on the materials obtained, are trying to work out numerical indices, which would constitute the minimum variability threshold of the estimates, the reaching of which would ensure the right to be a licence judge.