

Sexing animals using biometry: Intra-pair comparison is often superior to discriminant functions

Hanno Sandvik

Published on paper: 2001.
Published online: 2024-10-11.
ISSN 1502-4873 (paper).
ISSN 1891-5396 (electronic).
doi: <https://doi.org/10.5324/fn.v21i0.5993>.

Sandvik, H. 2001. Sexing animals using biometry: Intra-pair comparison is often superior to discriminant functions. - *Fauna norvegica* 21: 11-16.

In sexually monomorphic animal species it is difficult to sex unknown individuals. The most widely used method of sexing in those cases is discriminant analysis. In species that are socially monogamous and have biparental care, however, there exists another method which utilises the intra-pair difference in the trait used in sexing. Equations for calculating the fraction of wrongly sexed individuals using both methods are given and used to show that the latter method is more exact than discriminant analysis. The advantages and limits of both methods are compared. Field data from the kittiwake (*Rissa tridactyla*) are used to exemplify the intra-pair comparison method, and to test the assumptions underlying this method for this species. Among other things, it is shown that kittiwakes display random mating with respect to head+bill length. Sexing kittiwakes at Hornøya, northern Norway, by intra-pair differences in head+bill length gives correct results in 96% of the cases.

Key words: Sexing, biometry, discriminant function, random mating, *Rissa tridactyla*.

Hanno Sandvik, Department of Biology, University of Tromsø, 9037 Tromsø, Norway
(address for correspondence):

Norwegian Institute for Nature Research (NINA), Polar Environmental Centre, 9296 Tromsø, Norway
E-mail: hanno@ibg.uit.no

INTRODUCTION

In many animal species, sexes are indistinguishable by morphological characters. These sexually monomorphic species represent a problem for studies which need to sex the individuals investigated without dissecting them. A common solution to this problem is the calculation of discriminant functions using biometry (Sokal & Rohlf 1995): one or more biometric measures are taken in the field, which are afterwards put into discriminant functions published for the respective species. Depending on whether the variable resulting is above or below zero, animals are assigned to one or the other sex.

Many studies have referred to discriminant analyses as the method used in sexing animals. However, many of those studies, viz. the ones in which the mates of the individuals measured were known, could have used an alternative method. This method (hereafter referred to as intra-pair comparison) consists simply of assigning the larger individual of any pair to be a member of the larger sex and vice versa. The aim of this paper is to compare the accuracy of these two methods.

This is accomplished using computer simulation. Furthermore an example of the method is given using field data of the kittiwake

(*Rissa tridactyla*). Gulls and terns (Laridae) are an example of a sexually monomorphic taxon. Although this means that the sexes are indistinguishable by plumage, males are often slightly larger than females (see e.g., Croxall 1995, Székely et al. 2000), so that biometric characters such as head+bill length can be used in sexing (Coulson et al. 1983). For the kittiwake, Barrett et al. (1985) have described a discriminant function for distinguishing sexes by head+bill length. Using this method, they reported that 87% of kittiwakes at Hornøya in northern Norway could be sexed correctly. In this paper, I calculate the proportion of correctly sexed kittiwakes using intra-pair comparison.

MATERIAL AND METHODS

Assuming (1) random mating with respect to the trait used in sexing and (2) a normal distribution of this trait in both sexes, the fraction f_{pair} of wrongly sexed individuals using intra-pair comparison can be calculated by summing for all size classes of the larger sex the proportions of members that are mated with individuals larger than themselves:

$$f_{\text{pair}} = \sum_{i=0}^{\infty} \left[G(\mu_l, \sigma_l, (i-0.5) \cdot \Delta x) \cdot \Delta x \cdot \sum_{j=i}^{\infty} \left[G(\mu_s, \sigma_s, (j-0.5) \cdot \Delta x) \cdot \Delta x \right] \right] \quad (\text{Equation 1})$$

The fraction f_{discr} of wrongly sexed individuals using a discriminant function, presupposing the second of the above assumptions (normal distribution), is obtained by averaging the proportion of individuals of the larger sex that are smaller than the discriminant value and the proportion of individuals of the smaller sex that are larger than the discriminant value:

$$f_{\text{discr}} = \frac{1}{2} \sum_{i=0}^{x_{\text{discr}}/\Delta x} \left[G(\mu_l, \sigma_l, (i-0.5) \cdot \Delta x) \cdot \Delta x \right] + \frac{1}{2} \sum_{j=x_{\text{discr}}/\Delta x}^{\infty} \left[G(\mu_s, \sigma_s, (j-0.5) \cdot \Delta x) \cdot \Delta x \right] \quad (\text{Equation 2})$$

In both equations, i and j are integer variables; μ and σ are the mean and standard deviation of the trait used in sexing, subscripts (l and s) indicate the larger and the smaller sex; Δx is the breadth of the intervals used in the calculation; and G is the function of the Gaussian distribution:

$$G(\mu, \sigma, x) = \frac{e^{-\frac{(x-\mu)^2}{4\sigma^2}}}{\sigma \sqrt{2\pi}} \quad (\text{Equation 3})$$

x_{discr} denotes the discriminant value in Equation 2. If more than one character is used in sexing, x is simply replaced by the score of the discriminant function ($\sum \lambda_i x_i$; Sokal & Rohlf 1995).

For very small Δx , what Equation 1 does is to calculate the proportion of pairs in which the usual size relationship is reversed. Equation 2, on the other hand, calculates the proportion of individuals lying “on the wrong side” of x_{discr} .

The finite breadth of the intervals Δx accounts for the fact that no variable is known with infinite precision, i.e. for observer-, weather-related measuring errors, reading-off accuracy etc. Equations 1 and 2 are conservative in assuming that these errors always lower the precision of the sexing method (i.e., that measurements of individuals of the larger sex are always too small, and vice versa).

The precision of sexing by intra-pair comparison and by discriminant function was simulated by calculating f_{pair} and f_{discr} for a number of values of $\mu_l - \mu_s$, assuming $\sigma_l = \sigma_s$ and $\Delta x = 1\%$ of μ_s .

To test the two assumptions underlying Equation 1 for one species, I collected head+bill lengths from a total of 252 kittiwake breeding pairs during the breeding seasons 1995 and 1996 on the island of Hornøya (70° 22' N, 31° 10' E), northern Norway. Birds were caught using noose-poles. Head+bill length was measured with specially adapted callipers as described by Coulson et al. (1983) and Barrett et al. (1985). The head+bill length could be read off with an accuracy of 0.5 mm. However, the repeatability of measurements is a quite different matter from reading-off accuracy (Barrett et al. 1989). I therefore checked gauge repeatability

of head+bill lengths by comparing the two measurements of the 94 birds that had been caught and measured in both years.

Because the birds used were part of a demographic research project and killing them would thus have been detrimental to this project, they could not be sexed using dissection. Instead, birds were sexed using intra-pair comparison of head+bill length, assigning the largest bird to be male. This seemingly introduces a circular argument, but can actually only bias the results in two ways. First, the sexing method may somewhat underestimate the proportions of small males and large females, and thus slightly bias both means and variances of head+bill lengths in both sexes. However, this should not in itself lead to distributions closer to normality within the assumed sexes than within the real sexes. Second, the sexing method might make the mating pattern inferred less random (or, to be precise, more disassortative) than the real pattern. To investigate this possibility, I repeated the calculation of the regression of head+bill lengths (see below), constraining the sample to pairs in which the identity of the male was beyond reasonable doubt (i.e. head+bill length of one bird >93mm).

In 56 of the 131 breeding pairs captured in 1995, the birds retained their mates and were caught in the subsequent year as well. 26 of the 1995 breeding pairs divorced, i.e. at least one of the birds was caught breeding with a new mate in 1996. In order to obtain statistically independent observations (cf. Machlis et al. 1985), I calculated the mean of both measurements for all birds measured twice. If a bird had different mates in 1995 and 1996, both of these mates were also represented by a single value, viz. the mean of their head+bill lengths. This procedure resulted in a sample size of $n=173$ observations. As it is probably also defensible to use pairings as statistically independent observations, i.e. to count those birds twice that had different mates in both years, I repeated the calculation of the regression of head+bill lengths (see below) for this enlarged sample ($n=196$).

The assumption of random mating can be tested (a) by plotting head+bill lengths of the female against the ones of their mates; or (b) by analysing the variance of the differences in head+bill lengths between mates. The rationale of the latter method is

that assortative mating narrows down (and disassortative mating expands) the range of within-pair differences in head+bill lengths, compared to random mating. To carry out this method, observations from the pool of the 173 statistically independent data points were re-ordered, mimicking either perfectly assortative mating, perfectly disassortative mating, or random mating. While one re-ordering of measurements was enough to simulate assortative and disassortative mating (mating the smallest female with the smallest/largest male, respectively, etc.), I re-sampled several times without replacement ($n=12\ 000$) in order to obtain a reliable simulation of random mating, randomising "mate choice" in each replicate using a random generator. This method creates a null distribution, which also allows to give the expected standard deviation of the variance of the differences in head+bill lengths between mated kittiwakes.

The assumption of normally distributed head+bill lengths can be tested by simple inspection of a frequency plot, or by using appropriate statistical tests. I used the Shapiro-Wilk statistic W , calculated in PROC UNIVARIATE of the SAS software package (SAS Institute 1996).

All measures are given as means \pm 1 standard deviation.

RESULTS

Comparison of the two methods

The proportion of correctly sexed individuals was larger for intra-pair comparison than for discriminant analysis (Figure 1). This result held irrespective of the magnitude of the difference between the means of the sexes in the trait used in sexing, though the proportions approached each other for cases in which the means of each sex were either identical or very distinct. The difference between both methods of sexing was greatest when the mean difference between the sexes in the trait used in sexing equalled 1.7 standard deviations of the trait ($f_{\text{pair}}=12\%$, $f_{\text{discr}}=23\%$; cf. Figure 1).

Case study on kittiwakes from Hornøya

The head+bill lengths of kittiwakes obtained are given in Table 1. The mean difference between two measurements of the same bird in different years was 0.9 mm (median 1.0 mm, range 0.0–3.5 mm, 95% less than 2.5 mm, 90% less than 2.0 mm; $n=94$). Inserting these variables into Equation 1 with $\Delta x=1$ mm (the

Figure 1

The proportion of correctly sexed animals using biometry is a function of the magnitude of the difference between the sexes in the character used in sexing. Under the conditions regarded here, sexing by intra-pair comparison of individuals is superior to sexing by discriminant analysis irrespective of the difference between the sexes. The figure is derived from computer simulations using Equations 1 and 2, and assuming that the standard deviations are equal in both sexes and that the measuring inaccuracy is 1% of the mean of the smaller sex. See text for assumptions and limits of both methods.

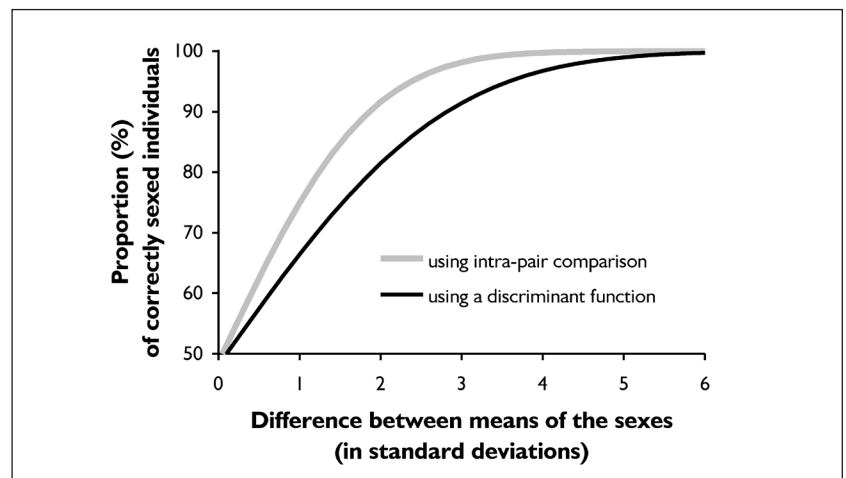


Table 1. Head+bill lengths of mated kittiwakes, captured at Hornøya, east Finnmark, in 1995 and 1996. The birds were sexed using intra-pair comparison (see text). For birds measured in both years, "individual birds" refers to the mean of both measurements. Measurements of "independent data points" do not necessarily refer to individual birds, but in 19 cases to means of two birds paired with the same mate in different years (see Material and Methods).

Variable	Sample size	Mean (mm)	Standard deviation (mm)
Female head+bill length (individual birds)	185	88.9	1.6
Male head+bill length (individual birds)	179	94.0	1.9
Female head+bill length (independent data points)	173	88.8	1.5
Male head+bill length (independent data points)	173	94.0	1.8

median gauge repeatability), the resulting fractions of wrongly sexed kittiwakes were $f_{\text{pair}}=3\%$ and $f_{\text{discr}}=8\%$.

The slope of the regression line of the head+bill lengths of the assumed females against the ones of their mates was not significantly different from zero (Figure 2). This result held also when restricting the sample to pairs with undoubted males (slope -0.03 ± 0.08 , $t=-0.35$, $p>0.5$, $n=116$, $r^2=0.00$), or when assuming the 196 observed pairings to be statistically independent (slope $+0.04\pm 0.09$, $t=0.49$, $p>0.5$, $r^2=0.00$).

The variances of the intra-pair differences in head+bill length obtained by re-ordering the real data set mimicking assortative

(disassortative) mating, were clearly lower (respectively, larger) than the observed one (Table 2). On the other hand, the observed variance did not depart significantly from the one obtained by random sampling (Table 2).

Measurements appeared to be more or less normally distributed in both sexes (Figure 3). The respective statistics showed that the distribution of head+bill lengths indeed did not differ significantly from normality in assumed females ($W=0.98$, $p>0.2$; skewness -0.07). In assumed males, however, the distribution was clearly (though not statistically significant at the 5% level: $W=0.97$, $p<0.1$) departing from normality, and was skewed towards larger values (skewness $+0.21$).

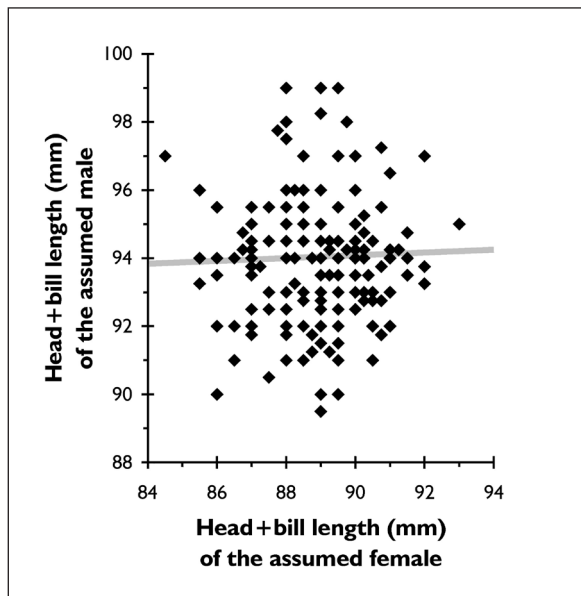


Figure 2

The relationship between head+bill lengths of paired kittiwakes breeding at Hornøya, east Finnmark, in 1995 and 1996. The slope of the regression line (light grey) is $+0.01\pm 0.10$ ($t=0.07$, $p>0.9$, $n=173$, $r^2=0.00$), indicating random mating with respect to head+bill length.

Table 2. Variances of within-pair differences in head+bill lengths of mated kittiwakes in relation to mating pattern. The real observation refers to kittiwakes captured at Hornøya, east Finnmark, in 1995 and 1996 (cf. Table 1, independent data points). This sample ($n=173$) was re-ordered, mimicking perfect assortative, disassortative and random mating (see Material and Methods for details). The observed variance departed less than 0.1 standard deviations from the variance expected under perfect random mating. (The mean difference in head+bill lengths within a pair was 5.2 mm in all cases.)

Mating pattern	Source	Constellations/iterations	Variance (mm ²)
Real	observation	1	5.66
Assortative	computer simulation	1	0.19
Disassortative	computer simulation	1	11.20
Random	computer simulation	12 000	5.69±0.43

DISCUSSION

Comparison of the two methods

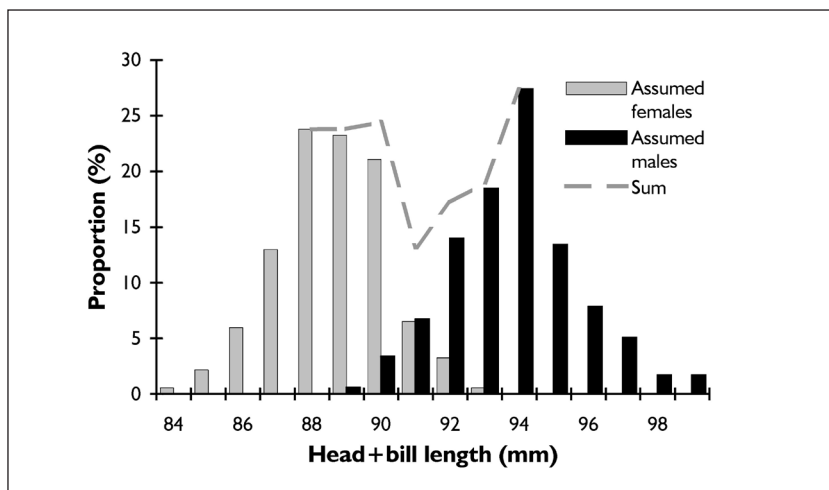
It can be seen from Figure 1 that sexing animals using intra-pair comparison is more reliable than using the more widely used method, viz. discriminant analysis. In cases where both sexes are very similar with respect to the trait used in sexing, both methods perform poorly. Likewise, when the character used in sexing approaches a disjunct distribution across sexes (right hand side of Figure 1), the difference between the methods diminishes.

This result is derived theoretically and is expected to hold whenever the assumptions underlying Equation 1 are met, i.e. the character used in sexing is normally distributed, and mating is random with respect to that character. Both assumptions are sensible null hypotheses for most species, and can be readily tested in cases of doubt. I performed one such test for the kittiwake, and was able to confirm both assumptions (see below).

Other researchers can use Equation 1 to calculate the error rate of this method of sexing for their respective species. Equation 1 is computationally easy to implement (a programme carrying out cal-

Figure 3

The frequency distribution of head+bill lengths of kittiwakes breeding at Hornøya, east Finnmark, in 1995 and 1996 ($n=185$ assumed females, 179 assumed males). The distributions inferred for both sexes are approximately normal, while the distribution of the total sample ("Sum", i.e. both sexes combined) is bimodal.



culations of both f_{pair} and f_{discr} can be obtained from the author), and does not require that only one character is used in sexing.

A further advantage of the intra-pair comparison method is that it is far more robust against regional size differences or changes in the size distribution of a given population over time. Discriminant functions reported in the literature do not necessarily apply to other populations of the same species (e.g. Barrett et al. 1985), and do not even need to be constant over time. The latter point is illustrated by a comparison of Barrett et al.'s results (1985: Table 6; females $89.2 \text{ mm} \pm 1.9 \text{ mm}$, $n=36$; males $94.5 \text{ mm} \pm 1.8 \text{ mm}$, $n=34$) with Table 1: my study reports smaller values for both sexes, the difference being as large as 1.5 mm in males, though both studies were performed in the same population. Such changes over time – in this case 15 years – clearly demonstrate the limits of discriminant functions.

Obviously, the animals have to have established pair-bonds if intra-pair comparison is to be applicable. This rules out those species that are not socially monogamous or do not have biparental care, but leaves for instance many taxa of birds. However, even in species with pair-bonds, exceptions do occur, especially when sex ratios are skewed. For a number of species in which one sex is overrepresented, female-female "pairs", male-male "pairs", triplets (e.g. female-female-male) etc. have been reported (see Hunt 1980, Nisbet & Hatch 1999 and references therein). There is, therefore, one case in which discriminant functions are superior to intra-pair comparisons even when the species has biparental care, viz. when the frequency of female-female "pairs" is higher than two times the difference between f_{pair} and f_{discr} . Male-male "pairs", being unable to produce eggs, do not represent any problem because they can be readily identified.

Case study on kittiwakes from Hornøya

Barrett et al. (1985) calculated a discriminant function for sexing adult kittiwakes using head+bill length. The fraction of wrongly

sexed birds reported for the colony of Hornøya was 13% (which departs somewhat from the 8% calculated above, probably due to Barrett et al.'s small sample size). However, had their sample been birds with known mates, at least 60% of the birds that were sexed wrongly using the discriminant function, could have been sexed correctly using intra-pair comparison – while all pairs sexed correctly using the discriminant function are necessarily sexed correctly using intra-pair comparison, too.

The assumptions inherent in Equation 1 were met by the sample I investigated. First, Table 2, Figure 2, and the results from large males show that mating was indeed random with respect to head+bill length. Among Charadriiformes, assortative mating by an aspect of bill size has only been reported for common terns (*Sterna hirundo*; Coulter 1986), razorbills (*Alca torda*; Wagner 1999), and possibly from Atlantic puffins (*Fratercula arctica*; Corkhill 1972, cf. Wagner 1999). Its absence in kittiwakes is evident from Figure 2, because head+bill lengths of mates were not related. Assortative mating would have reduced the fraction of wrongly sexed birds even more. Only disassortative mating would have reduced the precision of the intra-pair comparison method. However, for birds to mate disassortatively, the slope of the regression line would have to be negative in Figure 2, and the variance of within-pair differences in head+bill lengths would have to be larger than 5.69 mm^2 (see Table 2), both of which is clearly not the case in kittiwakes. Neither has it been reported from any other seabird.

Second, although the distribution of male head+bill lengths departed somewhat from normality, this does not underestimate the fraction of correctly sexed birds because the distribution was skewed to the right, and not to the left (Figure 3). That may mean that the proportion of males that can be mistaken for females was even smaller than if male head+bill lengths had been normally distributed. Alternatively, it could be a consequence of the method of sexing used (see Material and Methods). Furthermore, I found a clearly bimodal distribution of unsexed kittiwakes (hatched grey line in Figure 3) – in contrast to other studies (McGowan & Zonfrillo 1995) measuring kittiwakes outside the

breeding season. Their findings most likely reflect the mixing of birds of different localities during winter. This is a further reason why sexing mated kittiwakes is more reliable, because geographic variation in body size (e.g., Tatarinkova & Shklyarevich 1978, Barrett et al. 1985) has no bearing during the breeding season because of high degrees of philopatry in kittiwakes (but see Coulson & Nève de Mévergnies 1992).

Female-female "pairs" have been reported from the kittiwake (involving 2% of the pairings; Coulson & Thomas 1985) and from other Laridae, which tend to have female-biased sex ratios (Nisbet & Hatch 1999 and references therein). This lessens the proportion of correctly sexed individuals to approximately 96%. Several triplets have been observed in the sample investigated here ($n=12$). In all these cases it was possible to identify the male member, partly based on sexings of other years.

ACKNOWLEDGEMENTS

I would like to emphasise that I did not develop the method defended in this paper, I only justified it. I thank Rob T. Barrett, Per Fauchald and Magne Asheim for teaching me the method, and for helping to gather the data. The former, Tycho Anker-Nilssen, Thrine Moen Heggerget and an anonymous referee commented on the manuscript. The study was part of my Master thesis at the University of Tromsø, for which I received financial support from the Friedrich-Ebert-Stiftung.

SAMMENDRAG

Kjønnsbestemmelse av monogame dyr ved hjelp av biometriske mål bør baseres på sammenligninger innen par

Mange dyrearter kan bare kjønnsbestemmes ved disseksjon eller ved hjelp av biometriske målinger. Når man må kjønnsbestemme monogame dyr levende, er det best å basere seg på forskjeller innen par, dvs. simpelthen å anta at det største dyret i paret tilhører det største kjønn. Jeg viser hvordan man kan beregne feilprosenten ved kjønnsbestemmelse vha. både sammenligning innen par (Ligning 1) og diskriminantfunksjoner (Ligning 2). Ligning 1 bygger på to antakelser, som enkelt kan overprøves med felldata: tilfeldig pardannelse mht. karakteren som blir brukt i kjønnsbestemmelsen, og normalfordeling av denne karakteren i begge kjønn. Under de fleste betingelser gir den førstnevnte metoden en større nøyaktighet enn diskriminantfunksjoner (Figur 1). Bare i tilfeller der maken ikke er målt, eller ved hyppig forekomst av hunn-hunn-"par" vil diskriminantfunksjoner være overlegne.

Jeg illustrerer metoden med felldata av hekkende krykkjer på Hornøya i Øst-Finnmark (gitt i Tabell 1). Bare 4 % av fuglene vil kjønnsbestemmes feil vha. sammenligning av hode+nebb-

lengde innen par, hvilket er en 60 % lavere feilkvote enn det man oppnår ved bruk av den tilsvarende diskriminantfunksjonen.

Feltdataene støtter antakelsene bak Ligning 1 for krykkje: Pardannelse er tilfeldig mht. hode+nebb-lengde (bl.a. fordi hode+nebb-lengder ikke er relatert innen par, Figur 2; jf. Tabell 2), og hode+nebb-lengde er omtrentlig normalfordelt i begge kjønn (Figur 3).

REFERENCES

- Barrett, R.T., Fieler, R., Anker-Nilssen, T. & Rikardsen, F. 1985. Measurements and weight changes of Norwegian adult Puffins *Fratercula arctica* and Kittiwakes *Rissa tridactyla* during the breeding season. - Ringing Migr. 6: 102–112.
- Barrett, R.T., Peterz, M., Furness, R.W. & Durinck, J. 1989. The variability of biometric measurements. - Ringing Migr. 10: 13–16.
- Corkhill, P. 1972. Measurements of Puffins as criteria of sex and age. - Bird Study 19: 193–201.
- Coulson, J.C. & Nève de Mévergnies, G. 1992. Where do young Kittiwakes *Rissa tridactyla* breed, philopatry or dispersal? - Ardea 80: 187–197.
- Coulson, J.C. & Thomas, C.S. 1985. Changes in the biology of the kittiwake *Rissa tridactyla*: a 31-year study of a breeding colony. - J. Anim. Ecol. 54: 9–26.
- Coulson, J.C., Thomas, C.S., Butterfield, J.E.L., Duncan, N., Monaghan, P. & Shedden, C. 1983. The use of head and bill length to sex live gulls Laridae. - Ibis 125: 549–557.
- Coulter, M.C. 1986. Assortative mating and sexual dimorphism in the Common Tern. - Wilson Bull. 98: 93–100.
- Croxall, J.P. 1995. Sexual size dimorphism in seabirds. - Oikos 73: 399–403.
- Hunt, G.L., Jr. 1980. Mate selection and mating systems in seabirds. - Pp. 113–151 in Burger, J., Olla, B.L. & Winn, H.E. (eds.). Behavior of marine animals: current perspectives in research. Vol. 4: marine birds. Plenum, New York.
- Machlis, L., Dodd, P.W.D. & Fentress, J.C. 1985. The pooling fallacy: problems arising when individuals contribute more than one observation to the data set. - Z. Tierpsychol. 68: 201–214.
- McGowan, R.Y. & Zonfrillo, B. 1995. Pitfalls in sexing Kittiwakes *Rissa tridactyla* on head + bill length. - Ringing Migr. 16: 124–126.
- Nisbet, I.C.T. & Hatch, J.J. 1999. Consequences of a female-biased sex-ratio in a socially monogamous bird: female-female pairs in the Roseate Tern *Sterna dougallii*. - Ibis 141: 307–320.
- SAS Institute. 1996. The SAS system for Windows, version 6.12. - SAS Institute, Cary.
- Sokal, R.R. & Rohlf, F.J. 1995. Biometry: the principles and practice of statistics in biological research. 3rd ed. - Freeman, New York.
- Székely, T., Reynolds, J.D. & Figuerola, J. 2000. Sexual size dimorphism in shorebirds, gulls and alcids: the influence of sexual and natural selection. - Evolution 54: 1404–1413.
- Tatarinkova, I.P. & Shklyarevich, F.N. 1978. Sex identification of gulls (Laridae) by their head length. - Zool. Zh. 57: 1446–1447 (in Russian).
- Wagner, R.H. 1999. Sexual size dimorphism and assortative mating in Razorbills (*Alca torda*). - Auk 116: 542–544.