

# SAS/STAT® 13.1 User's Guide The INBREED Procedure



This document is an individual chapter from SAS/STAT® 13.1 User's Guide.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2013. SAS/STAT® 13.1 User's Guide. Cary, NC: SAS Institute Inc.

Copyright © 2013, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

**For a hard-copy book**: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

**U.S. Government License Rights; Restricted Rights:** The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a) and DFAR 227.7202-4 and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414.

December 2013

SAS provides a complete selection of books and electronic products to help customers use SAS® software to its fullest potential. For more information about our offerings, visit **support.sas.com/bookstore** or call 1-800-727-3228.

 $SAS^{\textcircled{@}}$  and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. @ indicates USA registration.

Other brand and product names are trademarks of their respective companies.



# Gain Greater Insight into Your SAS® Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.





# Chapter 50

# The INBREED Procedure

#### **Contents**

Getting Started: INBREED Procedure       397         The Format of the Input Data Set       397         Performing the Analysis       397         Syntax: INBREED Procedure       398         PROC INBREED Statement       398         BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400		
The Format of the Input Data Set       397         Performing the Analysis       397         Syntax: INBREED Procedure       398         PROC INBREED Statement       398         BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	Overview: INBREED Procedure	3977
Performing the Analysis       397         Syntax: INBREED Procedure       398         PROC INBREED Statement       398         BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis with BY Groups       400	Getting Started: INBREED Procedure	3978
Syntax: INBREED Procedure       398         PROC INBREED Statement       398         BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	The Format of the Input Data Set	3978
PROC INBREED Statement       398         BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	Performing the Analysis	3979
BY Statement       398         CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Missing Values       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	Syntax: INBREED Procedure	3983
CLASS Statement       398         GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	PROC INBREED Statement	3983
GENDER Statement       398         MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	BY Statement	3985
MATINGS Statement       398         VAR Statement       398         Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	CLASS Statement	3985
VAR Statement       398         Petails: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         xamples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	GENDER Statement	3985
Details: INBREED Procedure       398         Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	MATINGS Statement	3986
Missing Values       398         DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	VAR Statement	3986
DATA= Data Set       398         Computational Details       398         OUTCOV= Data Set       399         Displayed Output       399         ODS Table Names       399         Examples: INBREED Procedure       399         Example 50.1: Monoecious Population Analysis       399         Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	Details: INBREED Procedure	3987
Computational Details 398  OUTCOV= Data Set 399  Displayed Output 399  ODS Table Names 399  Examples: INBREED Procedure 399  Example 50.1: Monoecious Population Analysis 399  Example 50.2: Pedigree Analysis 399  Example 50.3: Pedigree Analysis with BY Groups 400	Missing Values	3987
OUTCOV= Data Set	DATA= Data Set	3987
Displayed Output	Computational Details	3988
ODS Table Names	OUTCOV= Data Set	3994
Examples: INBREED Procedure	Displayed Output	3996
Example 50.1: Monoecious Population Analysis399Example 50.2: Pedigree Analysis399Example 50.3: Pedigree Analysis with BY Groups400	ODS Table Names	3996
Example 50.2: Pedigree Analysis       399         Example 50.3: Pedigree Analysis with BY Groups       400	Examples: INBREED Procedure	3997
Example 50.3: Pedigree Analysis with BY Groups	Example 50.1: Monoecious Population Analysis	3997
	Example 50.2: Pedigree Analysis	3999
References	Example 50.3: Pedigree Analysis with BY Groups	4001
	References	4002

### **Overview: INBREED Procedure**

The INBREED procedure calculates the covariance or inbreeding coefficients for a pedigree. PROC INBREED is unique in that it handles very large populations.

The INBREED procedure has two modes of operation. One mode carries out analysis on the assumption that all the individuals belong to the same generation. The other mode divides the population into nonoverlapping generations and analyzes each generation separately, assuming that the parents of individuals in the current generation are defined in the previous generation.

PROC INBREED also computes averages of the covariance or inbreeding coefficients within sex categories if the sex of individuals is known.

## **Getting Started: INBREED Procedure**

This section demonstrates how you can use the INBREED procedure to calculate the inbreeding or covariance coefficients for a pedigree, how you can control the analysis mode if the population consists of nonoverlapping generations, and how you can obtain averages within sex categories.

For you to use PROC INBREED effectively, your input data set must have a definite format. The following sections first introduce this format for a fictitious population and then demonstrate how you can analyze this population by using the INBREED procedure.

#### The Format of the Input Data Set

The SAS data set used as input to the INBREED procedure must contain an observation for each individual. Each observation must include one variable identifying the individual and two variables identifying the individual's parents. Optionally, an observation can contain a known covariance coefficient and a character variable defining the gender of the individual.

For example, consider the following data:

```
data Population;
  input Individual $ Parent1 $ Parent2 $
        Covariance Sex $ Generation;
  datalines;
Mark George Lisa
Kelly Scott Lisa
                       F
Mike George Amy
      Mark Kelly 0.50 .
David Mark Kelly .
                       M 2
Merle Mike Jane
                       F
     Mark Kelly 0.50 M 2
Jim
Mark Mike
            Kelly
```

It is important to order the pedigree observations so that individuals are defined before they are used as parents of other individuals. The family relationships between individuals cannot be ascertained correctly unless you observe this ordering. Also, older individuals must precede younger ones. For example, 'Mark' appears as the first parent of 'David' at observation 5; therefore, his observation needs to be defined prior to observation 5. Indeed, this is the case (see observation 1). Also, 'David' is older than 'Jim', whose observation appears after the observation for 'David', as is appropriate.

In populations with distinct, nonoverlapping generations, the older generation (parents) must precede the younger generation. For example, the individuals defined in Generation=1 appear as parents of individuals defined in Generation=2.

PROC INBREED produces warning messages when a parent cannot be found. For example, 'Jane' appears as the second parent of the individual 'Merle' even though there are no previous observations defining her own parents. If the population is treated as an overlapping population, that is, if the generation grouping is ignored, then the procedure inserts an observation for 'Jane' with missing parents just before the sixth observation, which defines 'Merle' as follows:

```
Jane . . . F 2
Merle Mike Jane F 2
```

However, if generation grouping is taken into consideration, then 'Jane' is defined as the last observation in Generation=1, as follows:

```
Mike George Amy . M 1 Jane . . F 1
```

In this latter case, however, the observation for 'Jane' is inserted after the computations are reported for the first generation. Therefore, she does not appear in the covariance/inbreeding matrix, even though her observation is used in computations for the second generation (see Figure 50.2).

If the data for an individual are duplicated, only the first occurrence of the data is used by the procedure, and a warning message is displayed to note the duplication. For example, individual 'Mark' is defined twice, at observations 1 and 8. If generation grouping is ignored, then this is an error and observation 8 is skipped. However, if the population is processed with respect to two distinct generations, then 'Mark' refers to two different individuals, one in Generation=1 and the other in Generation=2.

If a covariance is to be assigned between two individuals, then those individuals must be defined prior to the assignment observation. For example, a covariance of 0.50 can be assigned between 'Mark' and 'Kelly' since they are previously defined. Note that assignment statements must have different formats depending on whether the population is processed with respect to generations (see the section "DATA= Data Set" on page 3987 for further information). For example, while observation 4 is valid for nonoverlapping generations, it is invalid for a processing mode that ignores generation grouping. In this latter case, observation 7 indicates a valid assignment, and observation 4 is skipped.

The latest covariance specification between any given two individuals overrides the previous one between the same individuals.

## **Performing the Analysis**

To compute the covariance coefficients for the overlapping generation mode, use the following statements:

```
proc inbreed data=Population covar matrix init=0.25;
run;
```

Here, the DATA= option names the SAS data set to be analyzed, and the COVAR and MATRIX options tell the procedure to output the covariance coefficients matrix. If you omit the COVAR option, the inbreeding coefficients are output instead of the covariance coefficients.

Note that the PROC INBREED statement also contains the INIT= option. This option gives an initial covariance between any individual and unknown individuals. For example, the covariance between any individual and 'Jane' would be 0.25, since 'Jane' is unknown, except when 'Jane' appears as a parent (see Figure 50.4).

Figure 50.1 Analysis for an Overlapping Population

			ne INBREED				
		Cox	variance Co	efficients			
Individual	Parent1	Parent2	George	Lisa	Mark	Scott	Kell
George			1.1250	0.2500	0.6875	0.2500	0.250
Lisa			0.2500	1.1250	0.6875	0.2500	0.687
Mark	George	Lisa	0.6875	0.6875	1.1250	0.2500	0.500
Scott			0.2500	0.2500	0.2500	1.1250	0.687
Kelly	Scott	Lisa	0.2500	0.6875	0.5000	0.6875	1.125
Amy			0.2500	0.2500	0.2500	0.2500	0.250
Mike	George	Amy	0.6875	0.2500	0.4688	0.2500	0.250
David	Mark	Kelly	0.4688	0.6875	0.8125	0.4688	0.812
Jane			0.2500	0.2500	0.2500	0.2500	0.250
Merle	Mike	Jane	0.4688	0.2500	0.3594	0.2500	0.250
Jim	Mark	Kelly	0.4688	0.6875	0.8125	0.4688	0.812
		Cov	variance Co	efficients			
Individual	Parent1	Parent2	Amy	Mike	David	Jane	Merl
George			0.2500	0.6875	0.4688	0.2500	0.468
Lisa			0.2500	0.2500	0.6875	0.2500	0.250
Mark	George	Lisa	0.2500	0.4688	0.8125	0.2500	0.359
Scott			0.2500	0.2500	0.4688	0.2500	0.250
Kelly	Scott	Lisa	0.2500	0.2500	0.8125	0.2500	0.250
Amy			1.1250	0.6875	0.2500	0.2500	0.468
Mike	George	Amy	0.6875	1.1250	0.3594	0.2500	0.687
David	Mark	Kelly	0.2500	0.3594	1.2500	0.2500	0.304
Jane			0.2500	0.2500	0.2500	1.1250	0.687
Merle	Mike	Jane	0.4688	0.6875	0.3047	0.6875	1.125
Jim	Mark	Kelly	0.2500	0.3594	0.8125	0.2500	0.304
		Cov	variance Co	efficients			
		Individual	Parent1	Parent2	Jim		
		George			0.4688		
		Lisa			0.6875		
		Mark	George	Lisa	0.8125		
		Scott	222-90		0.4688		
		Kelly	Scott	Lisa	0.8125		
		Amy			0.2500		
		Mike	George	Amy	0.3594		
		David	Mark	Kelly	0.8125		
		Jane		<b>-</b> 3	0.2500		
		Merle	Mike	Jane	0.3047		
		Jim	Mark	Kelly	1.2500		

In the previous example, PROC INBREED treats the population as a single generation. However, you might want to process the population with respect to distinct, nonoverlapping generations. To accomplish this, you need to identify the generation variable in a CLASS statement, as shown by the following statements:

```
proc inbreed data=Population covar matrix init=0.25;
   class Generation;
run;
```

Note that, in this case, the covariance matrix is displayed separately for each generation (see Figure 50.5).

Figure 50.2 Analysis for a Nonoverlapping Population

		The INBR	EED Procedure	e				
Generation = 1								
Covariance Coefficients								
Individua	l Parent	l Parent	.2 <b>M</b> a:	rk	Kelly	Mike		
Mark	George	Lisa	1.12	50	0.5000	0.4688		
Kelly	Scott	Lisa	0.50	00	1.1250	0.2500		
Mike	George	Amy	0.46	38	0.2500	1.1250		
		The INBR	EED Procedure	e				
		Gene	eration = 2					
		Covarianc	e Coefficient	ts				
Individual	Parent1	Parent2	David	Merle	Jim	Mark		
David	Mark	Kelly	1.2500	0.3047	0.8125	0.5859		
Merle	Mike	Jane	0.3047	1.1250	0.3047	0.4688		
Jim	Mark	Kelly	0.8125	0.3047	1.2500	0.5859		
Mark	Mike	Kelly	0.5859	0.4688	0.5859	1.1250		
		Number of	Individuals	4				

You might also want to see covariance coefficient averages within sex categories. This is accomplished by indicating the variable defining the gender of individuals in a GENDER statement and by adding the AVERAGE option to the PROC INBREED statement. For example, the following statements produce the covariance coefficient averages shown in Figure 50.3:

```
proc inbreed data=Population covar average init=0.25;
  class Generation;
  gender Sex;
run;
```

Figure 50.3 Averages within Sex Categories for a Nonoverlapping Generation

The	INBREED Procedure	e	
	Generation = 1		
Averages of Covariance	e Coefficient Mat:	rix in Generation 1	
	On Diagonal	Below Diagonal	
Male X Male	1.1250	0.4688	
Male X Female		0.3750	
Female X Female	1.1250	0.0000	
Over Sex	1.1250	0.4063	
	r of Males		
Number	r of Females	1	
Number	r of Individuals	3	
The	INBREED Procedure	e	
	Generation = 2		
Averages of Covariance	e Coefficient Mat	rix in Generation 2	
	On Diagonal	Below Diagonal	
Male X Male	1.2083	0.6615	
Male X Female		0.3594	
Female X Female	1.1250	0.0000	
Over Sex	1.1875	0.5104	
Vh	a of Wales	3	
	r of Males		
	r of Females		
Number	r of Individuals	4	

## **Syntax: INBREED Procedure**

The following statements are available in the INBREED procedure:

```
PROC INBREED < options > ;
BY variables;
CLASS variable;
GENDER variable;
MATINGS individual-list1 / mate-list1 < , ..., individual-listn / mate-listn > ;
VAR variables;
```

The PROC INBREED statement is required. Items within angle brackets (< >) are optional. The syntax of each statement is described in the following sections.

#### **PROC INBREED Statement**

PROC INBREED < options > ;

The PROC INBREED statement invokes the INBREED procedure. Table 50.1 summarizes the options available in the PROC INBREED statement.

Table 50.1 PROC INBREED Statement Options

Option	Description							
Specify Data	Specify Data Sets							
DATA=	Names the SAS data set							
OUTCOV=	Names an output data set to contain the inbreeding coefficients							
Control Type	e of Coefficient							
COVAR	Specifies that all coefficients output consist of covariance coefficients							
SELFDIAG	Includes an individual's self-mating kinship coefficient							
Control Disp	played Tables							
<b>AVERAGE</b>	Produces a table of averages of coefficients							
IND	Displays the individuals' inbreeding coefficients							
MATRIX	Displays the inbreeding coefficient matrix							
<b>Specify Defa</b>	ult Covariance Value							
INIT=	Specifies the covariance value							
Suppress Ou	ıtput							
INDL	Displays individuals' coefficients for only the last generation							
MATRIXL	Displays coefficients for only the last generation							
NOPRINT	Suppresses the display of all output							

#### **AVERAGE**

#### Α

produces a table of averages of coefficients for each pedigree of offspring. The AVERAGE option is used together with the GENDER statement to average the inbreeding/covariance coefficients within sex categories.

#### **COVAR**

C

specifies that all coefficients output consist of covariance coefficients rather than inbreeding coefficients.

#### DATA=SAS-data-set

names the SAS data set to be used by PROC INBREED. If you omit the DATA= option, the most recently created SAS data set is used.

#### **IND**

I

displays the individuals' inbreeding coefficients (diagonal of the inbreeding coefficients matrix) for each pedigree of offspring.

If you also specify the COVAR option, the individuals' covariance coefficients (diagonal of the covariance coefficients matrix) are displayed.

#### **INDL**

displays individuals' coefficients for only the last generation of a multiparous population.

#### INIT=cov

specifies the covariance value *cov* if any of the parents are unknown; a value of 0 is assumed if you do not specify the INIT= option.

#### **MATRIX**

#### М

displays the inbreeding coefficient matrix for each pedigree of offspring.

If you also specify the COVAR option, the covariance matrices are displayed instead of inbreeding coefficients matrices.

#### **MATRIXL**

displays coefficients for only the last generation of a multiparous population.

#### **NOPRINT**

suppresses the display of all output. Note that this option temporarily disables the Output Delivery System (ODS). For more information on ODS, see Chapter 20, "Using the Output Delivery System."

#### OUTCOV=SAS-data-set

names an output data set to contain the inbreeding coefficients. When the COVAR option is also specified, covariance estimates are output to the OUTCOV= data set instead of inbreeding coefficients.

#### **SELFDIAG**

includes an individual's self-mating kinship coefficient instead of the individual's inbreeding coefficient on the diagonal of the matrix in the OUTCOV= data set when the COVAR option is not specified.

#### **BY Statement**

#### BY variables;

You can specify a BY statement with PROC INBREED to obtain separate analyses of observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement for the INBREED procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

For more information about BY-group processing, see the discussion in SAS Language Reference: Concepts. For more information about the DATASETS procedure, see the discussion in the Base SAS Procedures Guide.

#### **CLASS Statement**

#### CLASS variable;

To analyze the population within nonoverlapping generations, you must specify the variable that identifies generations in a CLASS statement. Values of the generation variable, called *generation numbers*, must be integers, but generations are assumed to occur in the order of their input in the input data set rather than in numerical order of the generation numbers. The name of an individual needs to be unique only within its generation.

When the MATRIXL option or the INDL option is specified, each generation requires a unique generation number in order for the specified option to work correctly. If generation numbers are not unique, all the generations with a generation number that is the same as the last generation's are output.

#### **GENDER Statement**

#### **GENDER** variable;

The GENDER statement specifies a variable that indicates the sex of the individuals. Values of the sex variable must be character beginning with 'M' or 'F', for male or female. The GENDER statement is needed only when you specify the AVERAGE option to average the inbreeding/covariance coefficients within sex categories or when you want to include a gender variable in the OUTCOV= data set.

PROC INBREED makes the following assumptions regarding the gender of individuals:

- The first parent is always assumed to be the male. See the section "VAR Statement" on page 3986.
- The second parent is always assumed to be the female. See the section "VAR Statement" on page 3986.
- If the gender of an individual is missing or invalid, this individual is assumed to be a female unless the population is overlapping and this individual appears as the first parent in a later observation.

Any contradictions to these rules are reported in the SAS log.

#### **MATINGS Statement**

**MATINGS** individual-list1 / mate-list1 < , ..., individual-listn / mate-listn > ;

You can specify the MATINGS statement with PROC INBREED to specify selected matings of individuals. Each individual given in individual-list is mated with each individual given in mate-list. You can write multiple mating specifications if you separate them by commas or asterisks. The procedure reports the inbreeding coefficients or covariances for each pair of mates. For example, you can use the following statement to specify the mating of an individual named 'David' with an individual named 'Jane':

matings david / jane;

#### **VAR Statement**

**VAR** individual parent1 parent2 < covariance > ;

The VAR statement specifies three or four variables: the first variable contains an individual's name, the second variable contains the name of the individual's first parent, and the third variable contains the name of the individual's second parent. An optional fourth variable assigns a known value to the covariance of the individual's first and second parents in the current generation.

The first three variables in the VAR statement can be either numeric or character; however, only the first 12 characters of a character variable are recognized by the procedure. The fourth variable, if specified, must be numeric.

If you omit the VAR statement, then the procedure uses the first three unaddressed variables as the names of the individual and its parents. (Unaddressed variables are those that are not referenced in any other PROC INBREED statement.) If the input data set contains an unaddressed fourth variable, then it becomes the covariance variable.

#### **Details: INBREED Procedure**

#### **Missing Values**

A missing value for a parent implies that the parent is unknown. Unknown parents are assumed to be unrelated and not inbred unless you specify the INIT= option.

When the value of the variable identifying the individual is missing, the observation is not added to the list of individuals. However, for a multiparous population, an observation with a missing individual is valid and is used for assigning covariances.

Missing covariance values are determined from the INIT=cov option, if specified. Observations with missing generation variables are excluded.

If the gender of an individual is missing, it is determined from the order in which it is listed on the first observation defining its progeny for an overlapping population. If it appears as the first parent, it is set to 'M'; otherwise, it is set to 'F'. When the gender of an individual cannot be determined, it is assigned a default value of 'F'.

#### **DATA= Data Set**

Each observation in the input data set should contain necessary information such as the identification of an individual and the first and second parents of an individual. In addition, if a CLASS statement is specified, each observation should contain the generation identification; and, if a GENDER statement is specified, each observation should contain the gender of an individual. Optionally, each observation might also contain the covariance between the first and the second parents. Depending on how many statements are specified with the procedure, there should be enough variables in the input data set containing this information.

If you omit the VAR statement, then the procedure uses the first three *unaddressed variables* in the input data set as the names of the individual and his or her parents. Unaddressed variables in the input data set are those variables that are not referenced by the procedure in any other statements, such as CLASS, GENDER, or BY statements. If the input data set contains an unaddressed fourth variable, then the procedure uses it as the covariance variable.

If the individuals given by the variables associated with the first and second parents are not in the population, they are added to the population. However, if they are in the population, they must be defined prior to the observation that gives their progeny.

When there is a CLASS statement, the functions of defining new individuals and assigning covariances must be separated. This is necessary because the parents of any given individual are defined in the previous generation, while covariances are assigned between individuals in the current generation.

Therefore, there could be two types of observations for a multiparous population:

• one to define new individuals in the current generation whose parents have been defined in the previous generation, as in the following, where the missing value is for the covariance variable:

```
Mark George Lisa . M 1
Kelly Scott Lisa . F 1
```

• one to assign covariances between two individuals in the current generation, as in the following, where the individual's name is missing, 'Mark' and 'Kelly' are in the current generation, and the covariance coefficient between these two individuals is 0.50:

```
Mark Kelly 0.50 . 1
```

Note that the observations defining individuals must precede the observation assigning a covariance value between them. For example, if a covariance is to be assigned between 'Mark' and 'Kelly', then both of them should be defined prior to the assignment observation.

#### **Computational Details**

This section describes the rules that the INBREED procedure uses to compute the covariance and inbreeding coefficients. Each computational rule is explained by an example referring to the fictitious population introduced in the section "Getting Started: INBREED Procedure" on page 3978.

#### **Coancestry (or Kinship Coefficient)**

To calculate the inbreeding coefficient and the covariance coefficients, use the degree of relationship by descent between the two parents, which is called *coancestry* or *kinship coefficient* (Falconer and Mackay 1996, p.85), or *coefficient of parentage* (Kempthorne 1957, p.73). Denote the coancestry between individuals X and Y by  $f_{XY}$ . For information on how to calculate the coancestries among a population, see the section "Calculation of Coancestry" on page 3989.

#### **Covariance Coefficient (or Coefficient of Relationship)**

The covariance coefficient between individuals X and Y is defined by

$$Cov(X, Y) = 2f_{XY}$$

where  $f_{XY}$  is the coancestry between X and Y. The covariance coefficient is sometimes called the *coefficient* of relationship or the theoretical correlation (Falconer and Mackay (1996, p.153); Crow and Kimura (1970, p.134)). If a covariance coefficient cannot be calculated from the individuals in the population, it is assigned to an initial value. The initial value is set to 0 if the INIT= option is not specified or to cov if INIT=cov. Therefore, the corresponding initial coancestry is set to 0 if the INIT= option is not specified or to  $\frac{1}{2}cov$  if INIT=cov.

#### **Inbreeding Coefficients**

The inbreeding coefficient of an individual is the probability that the pair of alleles carried by the gametes that produced it are identical by descent (Falconer and Mackay (1996, Chapter 5), Kempthorne (1957, Chapter 5)). For individual X, denote its inbreeding coefficient by  $F_X$ . The inbreeding coefficient of an individual is equal to the coancestry between its parents. For example, if X has parents A and B, then the inbreeding coefficient of X is

$$F_{\rm X} = f_{\rm AB}$$

#### **Calculation of Coancestry**

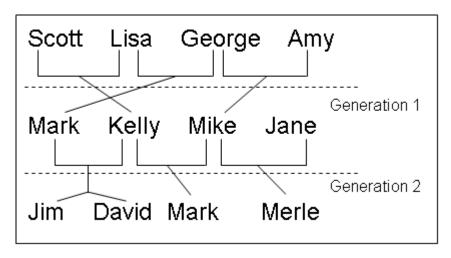
Given individuals X and Y, assume that X has parents A and B and that Y has parents C and D. For nonoverlapping generations, the basic rule to calculate the coancestry between X and Y is given by the following formula (Falconer and Mackay 1996, p.86):

$$f_{XY} = \frac{1}{4} (f_{AC} + f_{AD} + f_{BC} + f_{BD})$$

And the inbreeding coefficient for an offspring of X and Y, called Z, is the coancestry between X and Y:

$$F_{\rm Z} = f_{\rm XY}$$

Figure 50.4 Inbreeding Relationship for Nonoverlapping Population



For example, in Figure 50.4, 'Jim' and 'Mark' from Generation 2 are progenies of 'Mark' and 'Kelly' and of 'Mike' and 'Kelly' from Generation 1, respectively. The coancestry between 'Jim' and 'Mark' is

$$f_{\text{Jim,Mark}} = \frac{1}{\left(f_{\text{Mark,Mike}} + f_{\text{Mark,Kelly}} + f_{\text{Kelly,Mike}} + f_{\text{Kelly,Kelly}}\right)}$$

From the covariance matrix for Generation=1 in Figure 50.4 and the relationship that coancestry is half of the covariance coefficient,

$$f_{\text{Jim,Mark}} = \frac{1}{4} \left( \frac{0.4688}{2} + \frac{0.5}{2} + \frac{0.25}{2} + \frac{1.125}{2} \right) = 0.29298$$

For overlapping generations, if X is older than Y, then the basic rule can be simplified to

$$F_{\rm Z} = f_{\rm XY} = \frac{1}{2} (f_{\rm XC} + f_{\rm XD})$$

That is, the coancestry between X and Y is the average of coancestries between older X with younger Y's parents. For example, in Figure 50.5, the coancestry between 'Kelly' and 'David' is

$$f_{\text{Kelly,David}} = \frac{1}{2} \left( f_{\text{Kelly,Mark}} + f_{\text{Kelly,Kelly}} \right)$$

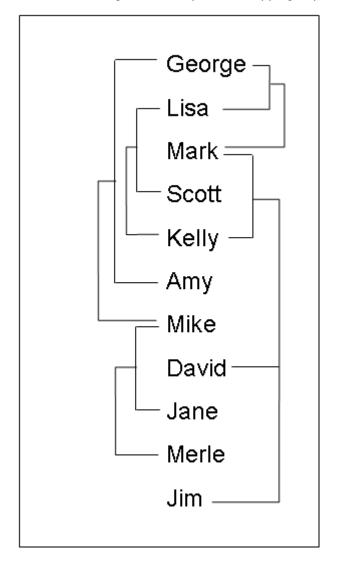


Figure 50.5 Inbreeding Relationship for Overlapping Population

This is so because 'Kelly' is defined before 'David'; therefore, 'Kelly' is not younger than 'David', and the parents of 'David' are 'Mark' and 'Kelly'. The covariance coefficient values Cov(Kelly,Mark) and Cov(Kelly,Kelly) from the matrix in Figure 50.5 yield that the coancestry between 'Kelly' and 'David' is

$$f_{\text{Kelly,David}} = \frac{1}{2} \left( \frac{0.5}{2} + \frac{1.125}{2} \right) = 0.40625$$

The numerical values for some initial coancestries must be known in order to use these rule. Either the parents of the first generation have to be unrelated, with f = 0 if the INIT= option is not specified in the PROC INBREED statement, or their coancestries must have an initial value of  $\frac{1}{2}cov$ , where cov is set by the INIT= option. Then the subsequent coancestries among their progenies and the inbreeding coefficients of their progenies in the rest of the generations are calculated by using these initial values.

Special rules need to be considered in the calculations of coancestries for the following cases.

#### Self-Mating

The coancestry for an individual X with itself,  $f_{XX}$ , is the inbreeding coefficient of a progeny that is produced by self-mating. The relationship between the inbreeding coefficient and the coancestry for self-mating is

$$f_{\rm XX} = \frac{1}{2} \left( 1 + F_{\rm X} \right)$$

The inbreeding coefficient  $F_X$  can be replaced by the coancestry between X's parents A and B,  $f_{AB}$ , if A and B are in the population:

$$f_{\rm XX} = \frac{1}{2} (1 + f_{\rm AB})$$

If X's parents are not in the population, then  $F_X$  is replaced by the initial value  $\frac{1}{2}cov$  if cov is set by the INIT= option, or  $F_X$  is replaced by 0 if the INIT= option is not specified. For example, the coancestry of 'Jim' with himself is

$$f_{\text{Jim,Jim}} = \frac{1}{2} \left( 1 + f_{\text{Mark,Kelly}} \right)$$

where 'Mark' and 'Kelly' are the parents of 'Jim'. Since the covariance coefficient Cov(Mark, Kelly) is 0.5 in Figure 50.5 and also in the covariance matrix for GENDER=1 in Figure 50.4, the coancestry of 'Jim' with himself is

$$f_{\text{Jim,Jim}} = \frac{1}{2} \left( 1 + \frac{0.5}{2} \right) = 0.625$$

When INIT=0.25, then the coancestry of 'Jane' with herself is

$$f_{\text{Jane,Jane}} = \frac{1}{2} \left( 1 + \frac{0.25}{2} \right) = 0.5625$$

because 'Jane' is not an offspring in the population.

#### Offspring and Parent Mating

Assuming that X's parents are A and B, the coancestry between X and A is

$$f_{XA} = \frac{1}{2} \left( f_{AB} + f_{AA} \right)$$

The inbreeding coefficient for an offspring of X and A, denoted by Z, is

$$F_{\rm Z} = f_{\rm XA} = \frac{1}{2} (f_{\rm AB} + f_{\rm AA})$$

For example, 'Mark' is an offspring of 'George' and 'Lisa', so the coancestry between 'Mark' and 'Lisa' is

$$f_{\text{Mark,Lisa}} = \frac{1}{2} \left( f_{\text{Lisa,George}} + f_{\text{Lisa,Lisa}} \right)$$

From the covariance coefficient matrix in Figure 50.5,  $f_{\text{Lisa,George}} = 0.25/2 = 0.125$ ,  $f_{\text{Lisa,Lisa}} = 1.125/2 = 0.5625$ , so that

$$f_{\text{Mark,Lisa}} = \frac{1}{2} (0.125 + 0.5625) = 0.34375$$

Thus, the inbreeding coefficient for an offspring of 'Mark' and 'Lisa' is 0.34375.

#### **Full Sibs Mating**

This is a special case for the basic rule given at the beginning of the section "Calculation of Coancestry" on page 3989. If X and Y are full sibs with same parents A and B, then the coancestry between X and Y is

$$f_{XY} = \frac{1}{4} (2f_{AB} + f_{AA} + f_{BB})$$

and the inbreeding coefficient for an offspring of A and B, denoted by Z, is

$$F_{\rm Z} = f_{\rm XY} = \frac{1}{4} (2f_{\rm AB} + f_{\rm AA} + f_{\rm BB})$$

For example, 'David' and 'Jim' are full sibs with parents 'Mark' and 'Kelly', so the coancestry between 'David' and 'Jim' is

$$f_{\text{David,Jim}} = \frac{1}{4} \left( 2 f_{\text{Mark,Kelly}} + f_{\text{Mark,Mark}} + f_{\text{Kelly,Kelly}} \right)$$

Since the coancestry is half of the covariance coefficient, from the covariance matrix in Figure 50.5,

$$f_{\text{David,Jim}} = \frac{1}{4} \left( 2 \times \frac{0.5}{2} + \frac{1.125}{2} + \frac{1.125}{2} \right) = 0.40625$$

#### Unknown or Missing Parents

When individuals or their parents are unknown in the population, their coancestries are assigned by the value  $\frac{1}{2}cov$  if cov is set by the INIT= option or by the value 0 if the INIT= option is not specified. That is, if either A or B is unknown, then

$$f_{\rm AB} = \frac{1}{2} cov$$

For example, 'Jane' is not in the population, and since 'Jane' is assumed to be defined just before the observation at which 'Jane' appears as a parent (that is, between observations 4 and 5), then 'Jane' is not older than 'Scott'. The coancestry between 'Jane' and 'Scott' is then obtained by using the simplified basic rule (see the section "Calculation of Coancestry" on page 3989):

$$f_{\text{Scott,Jane}} = \frac{1}{2} \left( f_{\text{Scott,}} + f_{\text{Scott,}} \right)$$

Here, dots (·) indicate Jane's unknown parents. Therefore,  $f_{Scott}$ , is replaced by  $\frac{1}{2}cov$ , where cov is set by the INIT= option. If INIT=0.25, then

$$f_{\text{Scott,Jane}} = \frac{1}{2} \left( \frac{0.25}{2} + \frac{0.25}{2} \right) = 0.125$$

For a more detailed discussion on the calculation of coancestries, inbreeding coefficients, and covariance coefficients, see Falconer and Mackay (1996); Kempthorne (1957); Crow and Kimura (1970).

#### OUTCOV= Data Set

The OUTCOV= data set has the following variables:

- a list of BY variables, if there is a BY statement
- the generation variable, if there is a CLASS statement
- the gender variable, if there is a GENDER statement
- \_Type\_, a variable indicating the type of observation. The valid values of the \_Type\_ variable are 'COV' for covariance estimates and 'INBREED' for inbreeding coefficients.
- Panel, a variable indicating the panel number used when populations delimited by BY groups contain different numbers of individuals. If there are n individuals in the first BY group and if any subsequent BY group contains a larger population, then its covariance/inbreeding matrix is divided into panels, with each panel containing n columns of data. If you put these panels side by side in increasing \_Panel\_ number order, then you can reconstruct the covariance or inbreeding matrix.

- \_Col\_, a variable used to name columns of the inbreeding or covariance matrix. The values of this variable start with 'COL', followed by a number indicating the column number. The names of the individuals corresponding to any given column *i* can be found by reading the individual's name across the row that has a \_Col\_ value of 'COL*i*'. When the inbreeding or covariance matrix is divided into panels, all the rows repeat for the first *n* columns, all the rows repeat for the next *n* columns, and so on.
- the variable containing the names of the individuals, that is, the first variable listed in the VAR statement
- the variable containing the names of the first parents, that is, the second variable listed in the VAR statement
- the variable containing the names of the second parents, that is, the third variable listed in the VAR statement
- a list of covariance variables Col1–Coln, where n is the maximum number of individuals in the first population

The functions of the variables \_Panel\_ and \_Col\_ can best be demonstrated by an example. Assume that there are three individuals in the first BY group and that, in the current BY group (Byvar=2), there are five individuals with the following covariance matrix.

COV	1	2	3	4	5
1	Cov(1,1)	Cov(1,2)	Cov(1,3)	Cov(1,4)	Cov(1,5)
2	Cov(2,1)	Cov(2,2)	Cov(2,3)	Cov(2,4)	Cov(2,5)
3	Cov(3,1)	Cov(3,2)	Cov(3,3)	Cov(3,4)	Cov(3,5)
4	Cov(4,1)	Cov(4,2)	Cov(4,3)	Cov(4,4)	Cov(4,5)
5	Cov(5,1)	Cov(5,2)	Cov(5,3)	Cov(5,4)	Cov(5,5)
Panel 1				Pan	el 2

Then the OUTCOV= data set appears as follows.

Byvar	_Panel_	_Col_	Individual	Parent	Parent2	Col1	Col2	Col3
2	1	COL1	1			Cov(1,1)	Cov(1,2)	Cov(1,3)
2	1	COL2	2			Cov(2,1)	Cov(2,2)	Cov(2,3)
2	1	COL3	3			Cov(3,1)	Cov(3,2)	Cov(3,3)
2	1		4			Cov(4,1)	Cov(4,2)	Cov(4,3)
2	1		5			Cov(5,1)	Cov(5,2)	Cov(5,3)
2	2		1			Cov(1,4)	Cov(1,5)	
2	2		2			Cov(2,4)	Cov(2,5)	
2	2		3			Cov(3,4)	Cov(3,5)	
2	2	COL1	4			Cov(4,4)	Cov(4,5)	
2	2	COL2	5			Cov(5,4)	Cov(5,5)	•

Notice that the first three columns go to the first panel (\_Panel\_=1), and the remaining two go to the second panel (\_Panel\_=2). Therefore, in the first panel, 'COL1', 'COL2', and 'COL3' correspond to individuals 1, 2, and 3, respectively, while in the second panel, 'COL1' and 'COL2' correspond to individuals 4 and 5, respectively.

#### **Displayed Output**

The INBREED procedure can output either covariance coefficients or inbreeding coefficients. Note that the following items can be produced for each generation if generations do not overlap.

The output produced by PROC INBREED can be any or all of the following items:

- a matrix of coefficients
- coefficients of the individuals
- · coefficients for selected matings

#### **ODS Table Names**

PROC INBREED assigns a name to each table it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in Table 50.2. For more information on ODS, see Chapter 20, "Using the Output Delivery System."

Table 50.2 ODS Tables Produced by PROC INBREED

ODS Table Name	Description	Statement	Option
AvgCovCoef	Averages of covariance coefficient matrix	GENDER	COVAR and AVERAGE
AvgInbreedingCoef	Averages of inbreeding coefficient matrix	GENDER	AVERAGE
CovarianceCoefficient	Covariance coefficient table	PROC	COVAR and MATRIX
InbreedingCoefficient	Inbreeding coefficient table	PROC	MATRIX
IndividualCovCoef	Covariance coefficients of individuals	PROC	IND and COVAR
IndividualInbreedingCoef	Inbreeding coefficients of individuals	PROC	IND
MatingCovCoef	Covariance coefficients of matings	MATINGS	COVAR
MatingInbreedingCoef	Inbreeding coefficients of matings	MATINGS	
NumberOfObservations	Number of observations	PROC	

# **Examples: INBREED Procedure**

#### **Example 50.1: Monoecious Population Analysis**

The following example shows a covariance analysis within nonoverlapping generations for a monoecious population. Parents of generation 1 are unknown and therefore assumed to be unrelated. The following statements produce Output 50.1.1 through Output 50.1.3:

```
data Monoecious;
   input Generation Individual Parent1 Parent2 Covariance @@;
  datalines;
              12...
11...
                             13..
2111 .
              2 2 1 2 .
                             2 3 2 3
3 1 1 2 .
              3 2 1 3 .
                             3 3 2 1
                             3 . 4 3 1.135
3 4 1 3 .
             3 . 2 3 0.50
title 'Inbreeding within Nonoverlapping Generations';
proc inbreed ind covar matrix data=Monoecious;
  class Generation;
run;
```

Output 50.1.1 Monoecious Population Analysis, Generation 1

```
Inbreeding within Nonoverlapping Generations
                       The INBREED Procedure
                           Generation = 1
                      Covariance Coefficients
Individual
              Parent1
                         Parent2
                                                                     3
                                       1.0000
1
2
                                                   1.0000
3
                                                               1.0000
               Covariance Coefficients of Individuals
          Individual
                        Parent1
                                   Parent2
                                                Coefficient
          1
                                                     1.0000
          2
                                                     1.0000
                                                     1.0000
                     Number of Individuals
```

Output 50.1.2 Monoecious Population Analysis, Generation 2

	Inbreeding	within None	overlapping (	Generations	
		The INBREE	) Procedure		
		Generat	tion = 2		
		Covariance (	Coefficients		
Individual	Parent1	Parent2	1	2	3
1	1	1	1.5000	0.5000	
2	1	2	0.5000	1.0000	0.2500
3	2	3	•	0.2500	1.0000
	Covaria	nce Coeffic	ients of Ind	ividuals	
	Individual	Parent1	Parent2	Coefficient	
	1	1	1	1.5000	
	2	1	2	1.0000	
	3	2	3	1.0000	
	N	umber of Ind	dividuals	3	

Output 50.1.3 Monoecious Population Analysis, Generation 3

Inbreeding within Nonoverlapping Generations  The INBREED Procedure									
		The Ir	BREED Proced	ure					
		Ge	eneration = 3						
		Covaria	ance Coefficie	ents					
Individual	Parent1	Parent2	1	2	3	4			
1	1	2	1.2500	0.5625	0.8750	0.5625			
2	1	3	0.5625	1.0000	1.1349	0.6250			
3	2	1	0.8750	1.1349	1.2500	1.1349			
4	1	3	0.5625	0.6250	1.1349	1.0000			
	Co	ovariance Coe	efficients of	Individuals	<b>;</b>				
	Individ	lual Parer	nt1 Parent2	2 Coeffi	cient				
	1	1	2	1	.2500				
	2 1 3 1.0000								
	3 2 1 1.2500								
	4	1	3	1	.0000				

#### Output 50.1.3 continued

```
Number of Individuals 4
```

Note that, since the parents of the first generation are unknown, off-diagonal elements of the covariance matrix are all 0s and on-diagonal elements are all 1s. If there is an INIT=cov value, then the off-diagonal elements would be equal to cov, while on-diagonal elements would be equal to 1 + cov/2.

In the third generation, individuals 2 and 4 are full siblings, so they belong to the same family. Since PROC INBREED computes covariance coefficients between families, the second and fourth columns of inbreeding coefficients are the same, except that their intersections with the second and fourth rows are reordered. Notice that, even though there is an observation to assign a covariance of 0.50 between individuals 2 and 3 in the third generation, the covariance between 2 and 3 is set to 1.135, the same value assigned between 4 and 3. This is because families get the same covariances, and later specifications override previous ones.

#### **Example 50.2: Pedigree Analysis**

Note the following from Output 50.2.1:

In the following example, an inbreeding analysis is performed for a complicated pedigree. This analysis includes computing selective matings of some individuals and inbreeding coefficients of all individuals. Also, inbreeding coefficients are averaged within sex categories. The following statements produce Output 50.2.1:

```
data Swine;
   input Swine_Number $ Sire $ Dam $ Sex $;
   datalines:
3504 2200 2501 M
3514 2521 3112
3519 2521 2501 F
2501 2200 3112 M
2789 3504 3514 F
3501 2521 3514 M
3712 3504 3514 F
3121 2200 3501 F
title 'Least Related Matings';
proc inbreed data=Swine ind average;
   var Swine_Number Sire Dam;
   matings 2501 / 3501 3504 ,
           3712 / 3121;
   gender Sex;
```

• Observation 4, which defines Swine\_Number=2501, should precede the first and third observations where the progeny for 2501 are given. PROC INBREED ignores observation 4 since it is given out of order. As a result, the parents of 2501 are missing or unknown.

• The first column in the "Inbreeding Averages" table corresponds to the averages taken over the ondiagonal elements of the inbreeding coefficients matrix, and the second column gives averages over the off-diagonal elements.

Output 50.2.1 Pedigree Analysis

		- Cargree		
	Leas	st Related Mati	ngs	
	The	INBREED Proced	lure	
	Inbreeding (	Coefficients of	Individuals	
Sw	rine_			
Nu	mber Sire	Dam	Coefficient	
22	100			
25	01		•	
35	04 2200	2501		
25	21			
31	.12			
35	2521	3112		
35	19 2521			
27	89 3504			
35	01 2521		0.2500	
	12 3504			
	.21 2200			
	Inbreeding	g Coefficients	of Matings	
	Sire	Dam	Coefficient	
	2501	3501		
		3504	0.2500	
	3712	3121	0.1563	
	Averages of I	nbreeding Coeff	icient Matrix	
	-	-		
		Inbreeding	Coancestr	У
Male	x Male	0.0625	0.104	2
Male	X Female		0.136	2
Fema	le X Female	0.0000	0.132	4
Over	Sex	0.0227	0.131	3
	Number	of Males	4	
		of Females	7	
		of Individuals		
	Number	or individuals	. 11	

#### **Example 50.3: Pedigree Analysis with BY Groups**

This example demonstrates the structure of the OUTCOV= data set created by PROC INBREED. Note that the first BY group has three individuals, while the second has five. Therefore, the covariance matrix for the second BY group is broken up into two panels. The following statements produce Output 50.3.1.

```
data Swine;
   input Group Swine_Number $ Sire $ Dam $ Sex $;
  datalines;
1 2789 3504 3514 F
2 2501 2200 3112
2 3504 2501 3782 M
proc inbreed data=Swine covar noprint outcov=Covariance
             init=0.4;
  var Swine_Number Sire Dam;
  gender Sex;
  by Group;
run;
title 'Printout of OUTCOV= data set';
proc print data=Covariance;
   format Col1-Col3 4.2;
run;
```

Output 50.3.1 Pedigree Analysis with BY Groups

					of OUT						
	_	_				Swine_		_			
Obs	Group	Sex	_TYPE_	_PANEL_	_COL_	Number	Sire	Dam	COL1	COL2	COL3
1	1	М	cov	1	COL1	3504			1.20	0.40	0.80
2	1	F	COV	1	COL2	3514			0.40	1.20	0.80
3	1	F	COV	1	COL3	2789	3504	3514	0.80	0.80	1.20
4	2	M	COV	1	COL1	2200			1.20	0.40	0.80
5	2	F	COV	1	COL2	3112			0.40	1.20	0.80
6	2	M	COV	1	COL3	2501	2200	3112	0.80	0.80	1.20
7	2	F	COV	1		3782			0.40	0.40	0.40
8	2	M	COV	1		3504	2501	3782	0.60	0.60	0.80
9	2	M	COV	2		2200			0.40	0.60	
10	2	F	COV	2		3112			0.40	0.60	
11	2	M	COV	2		2501	2200	3112	0.40	0.80	
12	2	F	COV	2	COL1	3782			1.20	0.80	
13	2	M	cov	2	COL2	3504	2501	3782	0.80	1.20	

# **References**

Crow, J. F. and Kimura, M. (1970), *An Introduction to Population Genetics Theory*, New York: Harper & Row.

Falconer, D. S. and Mackay, T. F. C. (1996), *Introduction to Quantitative Genetics*, 4th Edition, London: Longman.

Kempthorne, O. (1957), An Introduction to Genetic Statistics, New York: John Wiley & Sons.

# Subject Index

coefficient	population, overlapping, 3978, 3979, 3990
of relationship (INBREED), 3988	progeny, 3987, 3989, 3992, 3999
covariance coefficients, see INBREED procedure	second parent, 3986
	selective matings, output, 3986
full sibs mating	specifying gender, 3981
INBREED procedure, 3993	theoretical correlation, 3988
	unknown or missing parents, 3994
generation (INBREED)	variables, unaddressed, 3986, 3987
nonoverlapping, 3977, 3981	initial covariance value
number, 3985	assigning (INBREED), 3984
overlapping, 3977, 3979	INBREED procedure, 3987
variable, 3985	specifying (INBREED), 3979
INBREED procedure	mating
coancestry, computing, 3989	offspring and parent (INBREED), 3992, 3993
coefficient of relationship, computing, 3988	self (INBREED), 3992
covariance coefficients, 3977, 3979, 3981, 3984,	monoecious population analysis
3986, 3988	example (INBREED), 3997
covariance coefficients matrix, output, 3984	example (HABREED), 3777
first parent, 3986	offspring
full sibs mating, 3993	INBREED procedure, 3984, 3991
generation number, 3985	ordering observations
generation variable, 3985	INBREED procedure, 3978
generation, nonoverlapping, 3977, 3981	output data sets
generation, overlapping, 3977, 3979	OUTCOV= data set (INBREED), 3984, 3994
inbreeding coefficients, 3977, 3979, 3984, 3986,	output table names
3989	INBREED procedure, 3996
inbreeding coefficients matrix, output, 3984	1 ,
individuals, outputting coefficients, 3984	panels
individuals, specifying, 3981, 3985, 3986	INBREED procedure, 3994, 4001
initial covariance value, 3987	pedigree analysis
initial covariance value, assigning, 3984	example (INBREED), 3999, 4001
initial covariance value, specifying, 3979	INBREED procedure, 3977, 3978
kinship coefficient, 3988	population (INBREED)
last generation's coefficients, output, 3984	monoecious, 3997
mating, offspring and parent, 3992, 3993	multiparous, 3984, 3988
mating, self, 3992	nonoverlapping, 3985
matings, output, 3986	overlapping, 3978, 3979, 3990
monoecious population analysis, example, 3997	progeny
offspring, 3984, 3991	INBREED procedure, 3987, 3989, 3992, 3999
ordering observations, 3978	•
OUTCOV= data set, 3984, 3994	theoretical correlation
output table names, 3996	INBREED procedure, 3988
panels, 3994, 4001	
pedigree analysis, 3977, 3978	unknown or missing parents
pedigree analysis, example, 3999, 4001	INBREED procedure, 3994
population, monoecious, 3997	
population, multiparous, 3984, 3988	variables, unaddressed
population, nonoverlapping, 3985	INBREED procedure, 3986, 3987

# Syntax Index

AVERAGE option PROC INBREED statement, 3984
BY statement INBREED procedure, 3985
CLASS statement INBREED procedure, 3985 COVAR option
PROC INBREED statement, 3984
DATA= option PROC INBREED statement, 3984
GENDER statement, INBREED procedure, 3985
INBREED procedure syntax, 3983
INBREED procedure, BY statement, 3985
INBREED procedure, CLASS statement, 3985
INBREED procedure, GENDER statement, 3985
INBREED procedure, MATINGS statement, 3986
INBREED procedure, PROC INBREED statement, 3983
AVERAGE option, 3984
COVAR option, 3984
DATA= option, 3984
IND option, 3984
INDL option, 3984
INIT= option, 3984
MATRIX option, 3984
MATRIXL option, 3984
NOPRINT option, 3984
OUTCOV= option, 3984
SELFDIAG option, 3984
INBREED procedure, VAR statement, 3986
IND option
PROC INBREED statement, 3984
INDL option PROC INBREED statement, 3984
INIT= option
PROC INBREED statement, 3984
MATINGS statement, INBREED procedure, 3986 MATRIX option
PROC INBREED statement, 3984
MATRIXL option PROC INBREED statement, 3984

NOPRINT option

```
PROC INBREED statement, 3984

OUTCOV= option
PROC INBREED statement, 3984

PROC INBREED statement, see INBREED procedure

SELFDIAG option
PROC INBREED statement, 3984

VAR statement
INBREED procedure, 3986
```