

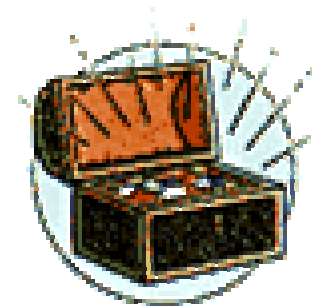
# Trésors cachés de la procédure FREQ



**Odds-ratios, risques  
relatifs, tests statistiques**

# Sommaire

1. **Les bases de la proc FREQ**
2. **Calculs supplémentaires**
  - Mesures d'association
  - Risques relatifs
  - Odds-ratios
3. **Tests statistiques**
  - Tests du chi-2
  - Tests de tendance
  - Test de proportion





# 1. Les bases de la proc FREQ

# Syntaxe (1/3)

Tableaux de  
fréquence  
simples

```
PROC FREQ DATA = tableSAS
```

```
  NLEVELS ;
```

cardinalités des variables  
citées dans TABLE

```
  TABLE variable1  
         variable2
```

```
  / NOPRINT NOPERCENT  
    NOFREQ NOCUM
```

n'édite pas une  
partie des sorties

```
  MISSING ;
```

considère les valeurs  
manquantes comme  
des valeurs normales

```
RUN ;
```

# Syntaxe (2/3)

Tableaux de  
fréquence  
croisés

```
PROC FREQ DATA = tableSAS ;
```

```
TABLE variable1* ← !
```

```
variable2 tableau non croisé
```

```
/ MISSING LIST SPARSE
```

```
NOPERCENT NOFREQ NOROW NOCOL ;
```

```
RUN ;
```

n'édite pas une  
partie des sorties

affiche tous les  
croisements en  
mode liste

# Syntaxe (3/3)

Tous types  
de tableaux

```
PROC FREQ DATA = tableSAS
ORDER = ...
;
TABLE ... ;
WEIGHT var ;
RUN ;
```

Variable de pondération  
(effectifs, sondage)

Ordre d'affichage

(INTERNAL par défaut) :

- FREQ (fréquences décroissantes)
- DATA (ordre dans la table lue)
- INTERNAL (valeurs sans format)
- FORMATTED (valeurs avec format)

# Objets ODS produits

Tableaux simples : `oneWayFreqs`

Tableaux croisés : `crossTabFreqs`

Tableaux croisés, option LIST : `list`

Cardinalités (option NLEVELS) : `nLevels`

Rappel de la syntaxe :

```
ODS OUTPUT nomObjet = tableSAS ;
```

# Une seule variable

```
PROC FREQ DATA = cours.color ;  
  TABLE eyes ;  
RUN ;
```

<b>Eye Color</b>				
<b>Eyes</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
<b>blue</b>	222	29.13	222	29.13
<b>brown</b>	341	44.75	563	73.88
<b>green</b>	199	26.12	762	100.00

# Deux variables (1/3)

Frequency  
Percent  
Row Pct  
Col Pct

```
PROC FREQ
DATA = cours.color ;
  TABLE hair * eyes ;
RUN ;
```

Hair(Hair Color)	Eyes(Eye Color)			Total
	blue	brown	green	
<b>black</b>	6 0.79 27.27 2.70	16 2.10 72.73 4.69	0 0.00 0.00 0.00	22 2.89
<b>dark</b>	51 6.69 28.02 22.97	94 12.34 51.65 27.57	37 4.86 20.33 18.59	182 23.88
<b>fair</b>	69 9.06 30.26 31.08	90 11.81 39.47 26.39	69 9.06 30.26 34.67	228 29.92
<b>medium</b>	68 8.92 31.34 30.63	94 12.34 43.32 27.57	55 7.22 25.35 27.64	217 28.48
<b>red</b>	28 3.67 24.78 12.61	47 6.17 41.59 13.78	38 4.99 33.63 19.10	113 14.83
<b>Total</b>	222 29.13	341 44.75	199 26.12	762 100.00

## Deux variables (2/3)

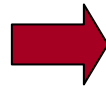
```
PROC FREQ DATA = cours.color ;
  TABLE hair * eyes / LIST ;
```

```
RUN ;
```

Hair	Eyes	Frequency	Percent	Cumulative Frequency	Cumulative Percent
black	blue	6	0.79	6	0.79
black	brown	16	2.10	22	2.89
dark	blue	51	6.69	73	9.58
dark	brown	94	12.34	167	21.92
dark	green	37	4.86	204	26.77
fair	blue	69	9.06	273	35.83
fair	brown	90	11.81	363	47.64
fair	green	69	9.06	432	56.69
medium	blue	68	8.92	500	65.62
medium	brown	94	12.34	594	77.95
medium	green	55	7.22	649	85.17
red	blue	28	3.67	677	88.85
red	brown	47	6.17	724	95.01
red	green	38	4.99	762	100.00

## Deux variables (3/3)

```
PROC FREQ DATA = cours.color ;
  TABLE hair * eyes / LIST SPARSE ;
RUN ;
```



Hair	Eyes	Frequency	Percent	Cumulative Frequency	Cumulative Percent
black	blue	6	0.79	6	0.79
black	brown	16	2.10	22	2.89
black	green	0	0.00	22	2.89
dark	blue	51	6.69	73	9.58
dark	brown	94	12.34	167	21.92
dark	green	37	4.86	204	26.77
fair	blue	69	9.06	273	35.83
fair	brown	90	11.81	363	47.64
fair	green	69	9.06	432	56.69
medium	blue	68	8.92	500	65.62
medium	brown	94	12.34	594	77.95
medium	green	55	7.22	649	85.17
red	blue	28	3.67	677	88.85
red	brown	47	6.17	724	95.01
red	green	38	4.99	762	100.00

# Cardinalités

```
PROC FREQ DATA = cours.color NLEVELS ;  
  TABLE hair eyes / NOPRINT ;  
RUN ;
```

Pas de tableaux  
de fréquences

**SAS v9  
seulement**

Number of Variable Levels		
Variable	Label	Levels
<b>Hair</b>	Hair Color	5
<b>Eyes</b>	Eye Color	3



## 2. Calculs supplémentaires

## Mesures d'association (1/4)

```
PROC FREQ DATA = ... ;  
    TABLE variable1 * variable2  
        / MEASURES ;  
  
RUN ;
```

L'option MEASURES demande l'édition de statistiques résumant la liaison entre les variables citées dans TABLE ; elles sont considérées comme ordinales.

## Mesures d'association (2/4)

```
PROC FREQ DATA = cours.jobSatisf ORDER = DATA ;
  TABLE income * satisf ;
RUN ;
```



**Table of income by satisf**

income	satisf				Total
	very dissatisfied	little dissatisfied	moderately satisfied	very satisfied	
<6000\$	20	24	80	82	206
6000-15000\$	22	38	104	125	289
15000-25000\$	13	28	81	113	235
>25000\$	7	18	54	92	171
<b>Total</b>	62	108	319	412	901

# Mesures d'association

## (3/4)

```
PROC FREQ DATA = cours.jobSatisf
      ORDER = DATA ;
      TABLE income * satisf
      / MEASURES  $\rightarrow$ 
      TEST GAMMA ;
RUN ;
```

<b>Test of H0: Gamma = 0</b>	
<b>ASE under H0</b>	0.0409
<b>Z</b>	3.0908
<b>One-sided Pr &gt; Z</b>	0.0010
<b>Two-sided Pr &gt;  Z </b>	0.0020

<b>Statistic</b>	<b>Value</b>	<b>ASE</b>
<b>Gamma</b>	0.1265	0.0408
<b>Kendall's Tau-b</b>	0.0876	0.0283
<b>Stuart's Tau-c</b>	0.0808	0.0261
<b>Somers' D C R</b>	0.0818	0.0265
<b>Somers' D R C</b>	0.0938	0.0303
<b>Pearson Correlation</b>	0.1030	0.0325
<b>Spearman Correlation</b>	0.1018	0.0329
<b>Lambda Asymmetric C R</b>	0.0000	0.0000
<b>Lambda Asymmetric R C</b>	0.0000	0.0000
<b>Lambda Symmetric</b>	0.0000	0.0000
<b>Uncertainty Coefficient C R</b>	0.0057	0.0033
<b>Uncertainty Coefficient R C</b>	0.0049	0.0028
<b>Uncertainty Coefficient Symmetric</b>	0.0053	0.0030

# Mesures d'association (4/4)

Calculées à partir du nombre de paires concordantes, discordantes, liées.

**GAMMA** : écart relatif du nombre de paires concordantes et discordantes ; se comporte comme une corrélation  $-1 \leq \gamma \leq 1$

**TAU-B DE KENDALL** : cf. Gamma, avec une correction sur le nombre de paires liées

**D DE SOMER** : différence de proportion entre paires concordantes et discordantes ; asymétrique (d'où 2 lignes colonnes/lignes et lignes/colonnes)

**CORRELATIONS** : calculées sur les rangs

## Risques relatifs (1/3)

```
PROC FREQ DATA = ... ;  
  TABLE variable1 * variable2  
        / RELRISK ;  
  
RUN ;
```

L'option RELRISK demande l'édition des risques relatifs pour les variables citées dans TABLE.

## Risques relatifs (2/3)

```
PROC FREQ DATA = cours.aspirine ORDER = DATA ;
  TABLE medic * attaque
    / RELRISK NOCOL NOPERCENT ;
  WEIGHT nb ;
RUN ;
```

**Table of medic by attaque**

medic	attaque		Total
	oui	non	
placebo	189 1.71	10845 98.29	11034
aspirine	104 0.94	10933 99.06	11037
<b>Total</b>	293	21778	22071

## Risques relatifs (3/3)

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	1.8321	1.4400	2.3308
Cohort (Col1 Risk)	1.8178	1.4330	2.3059
Cohort (Col2 Risk)	0.9922	0.9892	0.9953

Le risque relatif d'avoir une attaque en prenant un placebo plutôt que l'aspirine est de 1,82 (1,71 / 0,94).

Table of medic by attaque			
medic	attaque		Total
	oui	non	
placebo	189	10845	11034
	1.71	98.29	
aspirine	104	10933	11037
	0.94	99.06	
<b>Total</b>	293	21778	22071

## Odds-ratios (1/3)

```
PROC FREQ DATA = ... ;  
  TABLE variable1 * variable2  
        / RELRISK ;  
  
RUN ;
```

L'option RELRISK demande l'édition des odds-ratios pour les variables citées dans TABLE.

## Odds-ratios (2/3)

```
PROC FREQ DATA = cours.aspirine ORDER = DATA ;
  TABLE medic * attaque / RELRISK;
  WEIGHT nb ;
RUN ;
```

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	<b>1.8321</b>	1.4400	2.3308
Cohort (Col1 Risk)	<b>1.8178</b>	1.4330	2.3059
Cohort (Col2 Risk)	<b>0.9922</b>	0.9892	0.9953

La cote d'avoir une attaque est multipliée par 1,83 quand on prend un placebo plutôt que de l'aspirine.

# Odds-ratios (3/3)

medic	attaque		Total
	oui	non	
placebo	189 1.71	10845 98.29	11034
aspirine	104 0.94	10933 99.06	11037
<b>Total</b>	293	21778	22071

$$\text{OR} = (189 * 10933) / (104 * 10845) = 1,83$$

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	<b>1.8321</b>	1.4400	2.3308
Cohort (Col1 Risk)	<b>1.8178</b>	1.4330	2.3059
Cohort (Col2 Risk)	<b>0.9922</b>	0.9892	0.9953



## 3. Tests statistiques

## Chi-2 (1/3)

```
PROC FREQ DATA = ... ;  
    TABLE variable1 * variable2  
        / CHISQ CELLCHI2 EXPECTED ;  
RUN ;
```

L'option CHISQ demande l'édition des tests du chi-2 et des quantités associées. L'option CELLCHI2 permet d'afficher dans chaque case du tableau croisé sa contribution au chi-2. L'option EXPECTED permet d'afficher quel effectif théorique devrait peupler la case.

## Chi-2 (2/3)

```

PROC FREQ DATA = cours.aspirine ;
  TABLE medic * attaque
    / CHISQ CELLCHI2 EXPECTED
    NOROW NOCOL ;
  WEIGHT nb ;
RUN ;

```

Effectif théorique = produit des pourcentages marginaux et du nombre total d'observations

Contribution au Chi-2 =  $(\text{effectif théorique} - \text{effectif réel})^2 / (\text{effectif théorique})$

Frequency  
Expected  
Cell Chi-Square  
Percent

medic	attaque		Total
	oui	non	
placebo	189	10845	11034
	146.48	10888	
	12.343	0.1661	
	0.86	49.14	49.99
aspirine	104	10933	11037
	146.52	10890	
	12.339	0.166	
	0.47	49.54	50.01
Total	293	21778	22071
	1.33	98.67	100.00

## Chi-2 (3/3)

H0 : indépendance des deux variables du tableau croisé. Ici, l'indépendance semble très loin d'être vérifiée (p-value minuscule).

V de Cramer : rapport entre le chi-2 pour ce tableau et le plus fort chi-2 calculé pour un tableau de mêmes structure et effectif. Se lit comme un pourcentage d'une liaison parfaite entre les deux variables.

Statistic	DF	Value	Prob
Chi-Square	1	25.0139	<.0001
Likelihood Ratio Chi-Square	1	25.3720	<.0001
Continuity Adj. Chi-Square	1	24.4291	<.0001
Mantel-Haenszel Chi-Square	1	25.0128	<.0001
Phi Coefficient		0.0337	
Contingency Coefficient		0.0336	
Cramer's V		0.0337	

## Tendance (1/2)

```
PROC FREQ DATA = ... ;  
  TABLE variable1 * variable2  
        / TREND ;  
  
RUN ;
```

L'option TREND demande l'édition de tests de tendance.

## Tendance (2/2)

```
PROC FREQ DATA = cours.douleur ;
    WEIGHT count ;
    TABLE dose * adverse / TREND ;
RUN ;
```

### Cochran-Armitage Trend

<b>Statistic (Z)</b>	-4.7918
<b>One-sided Pr &lt; Z</b>	<.0001
<b>Two-sided Pr &gt;  Z </b>	<.0001

H0 = tendance linéaire du nombre de cas quand la variable en ligne (DOSE) augmente.

Test unilatéral : H1 = tendance à la baisse (car  $Z < 0$ )

Test bilatéral : H1 = tendance à la hausse ou à la baisse

# Test de proportion (1/3)

```
PROC FREQ DATA = ... ;  
    TABLE variableBinaire  
        / BINOMIAL  
        P=moyenne ;  
  
RUN ;
```

L'option `BINOMIAL` demande l'édition de tests. L'option `P` permet d'indiquer une valeur à tester pour la moyenne (proportion) de la variable citée dans `TABLE`.

# Test de proportion (2/3)

```
PROC FREQ DATA = cours.color ORDER=FREQ ;
  TABLE eyes / BINOMIAL NOCUM ;
RUN ;
```

Eye Color		
Eyes	Frequency	Percent
<b>brown</b>	341	44.75
<b>blue</b>	222	29.13
<b>green</b>	199	26.12

Binomial Proportion for Eyes = brown	
<b>Proportion</b>	0.4475
<b>ASE</b>	0.0180
<b>95% Lower Conf Limit</b>	0.4122
<b>95% Upper Conf Limit</b>	0.4828
<b>Exact Conf Limits</b>	
<b>95% Lower Conf Limit</b>	0.4118
<b>95% Upper Conf Limit</b>	0.4836

# Test de proportion (3/3)

<b>Test of H0: Proportion = 0.5</b>	
<b>ASE under H0</b>	0.0181
<b>Z</b>	-2.8981
<b>One-sided Pr &lt; Z</b>	0.0019
<b>Two-sided Pr &gt;  Z </b>	0.0038

← H1 : proportion < 0.5

← H1 : proportion ≠ 0.5

Sample Size = 762

L'hypothèse alternative du test unilatéral serait « proportion > 0.5 » si Z était positif.