

```
> tests <- single.snp.tests(cc, data = subject.support, snp.data = snps.10)
```

Some words of explanation are required. In the call, the `snp.data=` argument is mandatory and provides the name of the matrix providing the genotype data. The `data=` argument gives the name of the data frame that contains the remaining arguments — usually the subject support data frame⁵.

Let us now see what has been calculated:

```
> summary(tests)
```

| | N | Chi.squared.1.df | Chi.squared.2.df | P.1df |
|----------|---------|------------------|------------------|----------------|
| Min. | :974 | Min. : 0.0000 | Min. : 0.0000 | Min. :0.0000 |
| 1st Qu.: | :987 | 1st Qu.: 0.1724 | 1st Qu.: 0.7904 | 1st Qu.:0.1410 |
| Median : | :989 | Median : 0.7729 | Median : 1.8559 | Median :0.3793 |
| Mean : | :989 | Mean : 1.5608 | Mean : 2.5968 | Mean :0.4192 |
| 3rd Qu.: | :991 | 3rd Qu.: 2.1670 | 3rd Qu.: 3.6651 | 3rd Qu.:0.6780 |
| Max. : | :999 | Max. :34.0217 | Max. :37.2487 | Max. :1.0000 |
| | | NA's :4 | NA's :830 | NA's :4 |
| P.2df | | | | |
| Min. | :0.0000 | | | |
| 1st Qu.: | :0.1600 | | | |
| Median : | :0.3954 | | | |
| Mean : | :0.4282 | | | |
| 3rd Qu.: | :0.6736 | | | |
| Max. : | :1.0000 | | | |
| NA's : | :830 | | | |

We have, for each SNP, chi-squared tests on 1 and 2 degrees of freedom (df), together with N , the number of subjects for whom data were available. The 1 df test is the familiar Cochran-Armitage test for codominant effect while the 2 df test is the conventional Pearsonian test for the 3×2 contingency table. The large number of NA values for the latter test reflects the fact that, for these SNPs, the minor allele frequency was such that one homozygous genotype did not occur in the data.

We will probably wish to restrict our attention to SNPs that pass certain criteria. For example

```
> use <- snpsum$MAF > 0.01 & snpsum$z.HWE^2 < 200
```

⁵This is not mandatory — we could have made `cc` available in the global environment. However we would then have to be careful that the values are in the right order; by specifying the data frame, order is forced to be correct by checking the order of the row names for the `data` and `snp.data` arguments.

```
> tests <- single.snp.tests(cc, data = subject.support, snp.data = snps.10)
```

Some words of explanation are required. In the call, the `snp.data=` argument is mandatory and provides the name of the matrix providing the genotype data. The `data=` argument gives the name of the data frame that contains the remaining arguments — usually the subject support data frame⁵.

Let us now see what has been calculated:

```
> summary(tests)
```

| | N | Chi.squared.1.df | Chi.squared.2.df | P.1df |
|----------|---------|------------------|------------------|----------------|
| Min. | :974 | Min. : 0.0000 | Min. : 0.0000 | Min. :0.0000 |
| 1st Qu.: | :987 | 1st Qu.: 0.1724 | 1st Qu.: 0.7915 | 1st Qu.:0.1410 |
| Median : | :989 | Median : 0.7729 | Median : 1.8561 | Median :0.3793 |
| Mean : | :989 | Mean : 1.5608 | Mean : 2.5965 | Mean :0.4192 |
| 3rd Qu.: | :991 | 3rd Qu.: 2.1670 | 3rd Qu.: 3.6642 | 3rd Qu.:0.6780 |
| Max. : | :999 | Max. :34.0217 | Max. :37.2487 | Max. :1.0000 |
| | | NA's :4 | NA's :810 | NA's :4 |
| P.2df | | | | |
| Min. | :0.0000 | | | |
| 1st Qu.: | :0.1601 | | | |
| Median : | :0.3953 | | | |
| Mean : | :0.4281 | | | |
| 3rd Qu.: | :0.6732 | | | |
| Max. : | :1.0000 | | | |
| NA's : | :810 | | | |

We have, for each SNP, chi-squared tests on 1 and 2 degrees of freedom (df), together with N , the number of subjects for whom data were available. The 1 df test is the familiar Cochran-Armitage test for codominant effect while the 2 df test is the conventional Pearsonian test for the 3×2 contingency table. The large number of NA values for the latter test reflects the fact that, for these SNPs, the minor allele frequency was such that one homozygous genotype did not occur in the data.

We will probably wish to restrict our attention to SNPs that pass certain criteria. For example

```
> use <- snpsum$MAF > 0.01 & snpsum$z.HWE^2 < 200
```

⁵This is not mandatory — we could have made `cc` available in the global environment. However we would then have to be careful that the values are in the right order; by specifying the data frame, order is forced to be correct by checking the order of the row names for the `data` and `snp.data` arguments.