**Mean and Variance of a Dichotomous Variable (coded 0,1)**

The mean is the proportion of the total scores that are in the group coded 1 (*p*). The variance is *p* times *q*, where *p* is the proportion of total scores the in the one group and *q* is the proportion of scores in the other group. Of course, *p* + *q* = 1. Here is an empirical demonstration of this. I made the sample size very large to reduce the impact of SPSS computing the variances as the ratio of sum of squares divided by *N*-1 rather than divided by *N*.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .50 | .500 | .250 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .60 | .490 | .240 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .80 | .400 | .160 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .70 | .458 | .210 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .90 | .300 | .090 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 0 | 1 | .999 | .032 | .001 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
| Dichot | 1000000 | 1 | 1 | 1.00 | .000 | .000 |
| Valid N (listwise) | 1000000 |  |  |  |  |  |

As the two groups *n*’s get more and more unequal, the mean gets further and further away from .5 and variance get closer and closer to zero.