

(\* BIOEN 3070/6070: Introduction to Statistics for Bioengineers \*)

(\* © Orly Alter 2016 \*)

(\* In-Class Project 3: Chi-Squared Test of Goodness of Fit \*)

(\* The Wikipedia Example \*)

```
e = {50, 50};  
o = {44, 56};  
chi = N[Sum[(o[[a]] - e[[a]])^2 / e[[a]], {a, 1, Dimensions[o][[1]]}]]
```

1.44

? PDF

```
PDF[ChiSquareDistribution[1], chi]
```

---

PDF[*dist*, *x*] gives the probability density function for the symbolic distribution *dist* evaluated at *x*.  
PDF[*dist*, {*x*<sub>1</sub>, *x*<sub>2</sub>, ...}] gives the multivariate probability density function for a symbolic distribution *dist* evaluated at {*x*<sub>1</sub>, *x*<sub>2</sub>, ...}.  
PDF[*dist*] gives the PDF as a pure function. >>

0.161822

? CDF

```
CDF[ChiSquareDistribution[1], chi]  
pValue = 1 - CDF[ChiSquareDistribution[1], chi]  
If[pValue < 0.05, "Not Likely", "Not Not Likely"]
```

---

CDF[*dist*, *x*] gives the cumulative distribution function for the symbolic distribution *dist* evaluated at *x*.  
CDF[*dist*, {*x*<sub>1</sub>, *x*<sub>2</sub>, ...}] gives the multivariate cumulative distribution function for the symbolic distribution *dist* evaluated at {*x*<sub>1</sub>, *x*<sub>2</sub>, ...}.  
CDF[*dist*] gives the CDF as a pure function. >>

0.769861

0.230139

Not Not Likely

(\* Lobo (Nat Ed 2008) \*)

(\* Harris (1912) \*)

```
e = {300, 100};  
o = {305, 95};  
chi = N[Sum[(o[[a]] - e[[a]])^2 / e[[a]], {a, 1, Dimensions[o][[1]]}]]
```

0.333333

```
CDF[ChiSquareDistribution[1], chi]  
pValue = 1 - CDF[ChiSquareDistribution[1], chi]  
If[pValue < 0.05, "Not Likely", "Not Not Likely"]
```

0.436297

0.563703

Not Not Likely

(\* Variation on Harris (1912) \*)

```
e = {300, 100};  
o = {50, 350};  
chi = N[Sum[(o[[a]] - e[[a]])^2 / e[[a]], {a, 1, Dimensions[o][[1]]}]]  
833.333
```

```
pValue = 1 - CDF[ChiSquareDistribution[1], chi]  
If[pValue < 0.05, "Not Likely", "Not Not Likely"]  
0.
```

Not Likely

(\* Lobo and Shaw (Nat Ed 2008) \*)

(\* Bateson, Saunders and Punnett (1905) \*)

```
e = {1199, 400, 400, 133};  
o = {1528, 106, 117, 381};  
chi = N[Sum[(o[[a]] - e[[a]])^2 / e[[a]], {a, 1, Dimensions[o][[1]]}]]  
969.025
```

```
pValue = 1 - CDF[ChiSquareDistribution[3], 969.025]  
If[pValue < 0.05, "Not Likely", "Not Not Likely"]  
1.
```

0.

Not Likely