joint, marginal and conditional distributions

Contingency table based on relative frequencies

example

Suppose we are interested in the relationship between an individual's hair (X) and eye (Y) color.

	X				
Y	$P(X, Y)$	blonde	red	brown	black	$\sum$
	blue	0.12	0.05	0.12	0.01	0.30
	green	0.12	0.07	0.09	0	0.28
	brown	0.16	0.07	0.16	0.03	0.42
	$\sum$	0.40	0.19	0.37	0.04	1.00

Bernoulli random variable

$$P(X = x \mid p) = \begin{cases} p, & \text{if } x = 1 \\ 1 - p, & \text{if } x = 0 \end{cases}$$

- A random variable for modeling binary events
- Two possible outcomes:
  - Success - value 1
  - Failure - value 0

X	$P(X = x)$
0	$1 - p$
1	p

- Single parameter p, probability of a success
- multiple Bernoulli r.v. can be combined to model more complex random variables
- Shorthand notation: $X \sim \text{Bern}(p)$
- $E(X) = p,\ V(X) = p(1 - p)$