Neuroinformatics Lecture (L6)

Prof. Dr. Gordon Pipa
Institute of Cognitive Science, University of Osnabrück

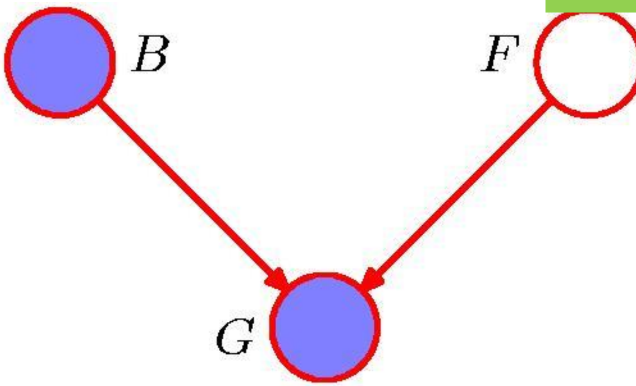
Rehearsal

R

$$p(F=0) = 0.1$$

$$p(F=0|G=0) \simeq 0.257$$

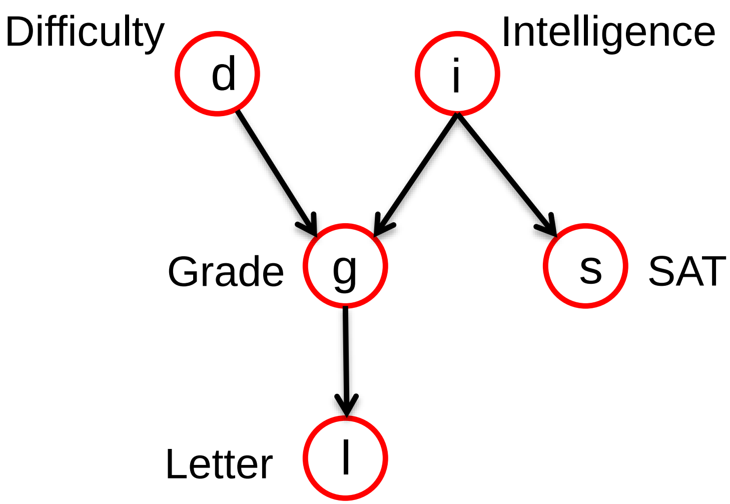
$$p(F=0|G=0,B=0) \simeq 0.111$$



Probability of an empty tank is changed by observing B = 0.

This is referred to as "explaining away":
Finding out that the battery is flat *explains away* the observation that the fuel gauge reads empty.
(In the example this leads to a reduction of the probability. For other example the probability may increase if we explain away another possible explanation)

An Example: Now with real p's



$$p(d,i,g,s,l) = p(d)\,p(i)\,p(g\,|\,i,d)\,p(s\,|\,i)\,p(l\,|\,g)$$

What is the probability of a good letter p(l=1)?

We marginalize the joint probability.

$$p(l=1) = \sum_{d,i,g,s} p(d,i,g,s,l)$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

$$p(l=1)=\sum_{d,i,g,s} p(d)p(i)p(g\mid i,d)p(s\mid i)p(l\mid g)$$

$$p(l=1)=$$
$$\boxed{0.6\cdot 0.7\cdot 0.3\cdot 0.95\cdot 0.9+}$$
$$0.4\cdot 0.7\cdot 0.05\cdot 0.95\cdot 0.9+$$
$$0.6\cdot 0.3\cdot 0.9\cdot 0.2\cdot 0.9+$$
$$0.4\cdot 0.3\cdot 0.5\cdot 0.2\cdot 0.9+\ldots.$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

$$p(l=1)=\sum_{d,i,g,s} p(d)p(i)p(g\,|\,i,d)p(s\,|\,i)p(l\,|\,g)$$

$$p(l=1)=$$

$$0.6\cdot 0.7\cdot 0.3\cdot 0.95\cdot 0.9+$$

$$\boxed{0.4\cdot 0.7\cdot 0.05\cdot 0.95\cdot 0.9+}$$

$$0.6\cdot 0.3\cdot 0.9\cdot 0.2\cdot 0.9+$$

$$0.4\cdot 0.3\cdot 0.5\cdot 0.2\cdot 0.9+....$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

$$p(l=1)=\sum_{d,i,g,s} p(d)p(i)p(g\mid i,d)p(s\mid i)p(l\mid g)$$

$$p(l=1)=$$

$$0.6\cdot 0.7\cdot 0.3\cdot 0.95\cdot 0.9+$$

$$0.4\cdot 0.7\cdot 0.05\cdot 0.95\cdot 0.9+$$

$$\boxed{0.6\cdot 0.3\cdot 0.9\cdot 0.2\cdot 0.9+}$$

$$0.4\cdot 0.3\cdot 0.5\cdot 0.2\cdot 0.9+\ldots.$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

$$p(l=1)=\sum_{d,i,g,s} p(d)p(i)p(g\,|\,i,d)p(s\,|\,i)p(l\,|\,g)$$

$$p(l=1)=$$
$$0.6\cdot0.7\cdot0.3\cdot0.95\cdot0.9+$$
$$0.4\cdot0.7\cdot0.05\cdot0.95\cdot0.9+$$
$$0.6\cdot0.3\cdot0.9\cdot0.2\cdot0.9+$$
$$\boxed{0.4\cdot0.3\cdot0.5\cdot0.2\cdot0.9+....}$$

Sum over all 2*2*3*2 =24 combinations

$$p(l=1)=0.502$$

The probability that you get a good letter without knowing anything about d,i,g,s is 0.502.

d=0	d=1
0.6	0.4

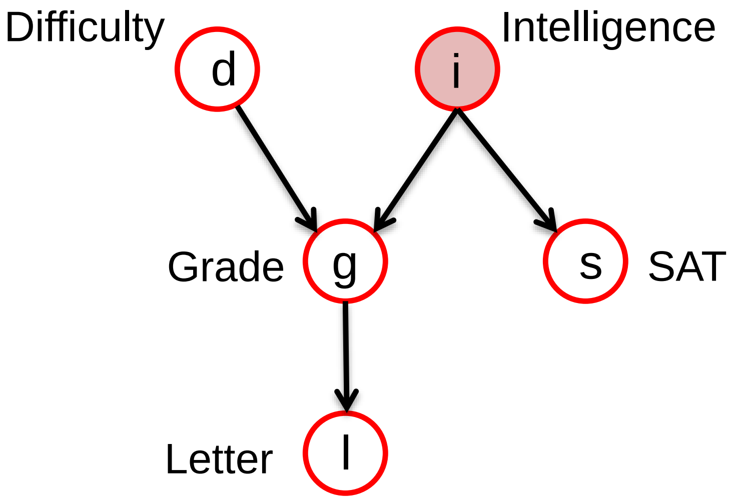
i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p’s



$$p(d,i,g,s,l) = p(d)\,p(i)\,p(g \mid i,d)\,p(s \mid i)\,p(l \mid g)$$

What is the probability of a good letter given that the student is normally intelligent

p(l=1|i=0)

$$p(d,i,g,s,l) = p(d,g,s,l \mid i)\,p(i) = p(d)\,p(i)\,p(g \mid i,d)\,p(s \mid i)\,p(l \mid g)$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

What is the probability of a good letter given that the student is not intelligent

p(l=1|i=0)

$$p(d,g,s,l\,|\,i) = p(d)\,p(g\,|\,i,d)\,p(s\,|\,i)\,p(l\,|\,g)$$

Marginalizing the conditional probability

$$p(l=1\,|\,i) = \sum_{\boxed{d,g,s}} p(d)\,p(g\,|\,i,d)\,p(s\,|\,i)\,p(l\,|\,g)$$

Sum over all $\boxed{2*3*2 = 12}$ combinations

$$p(l=1\,|\,i=0) = 0.389$$

d=0	d=1
0.6	0.4

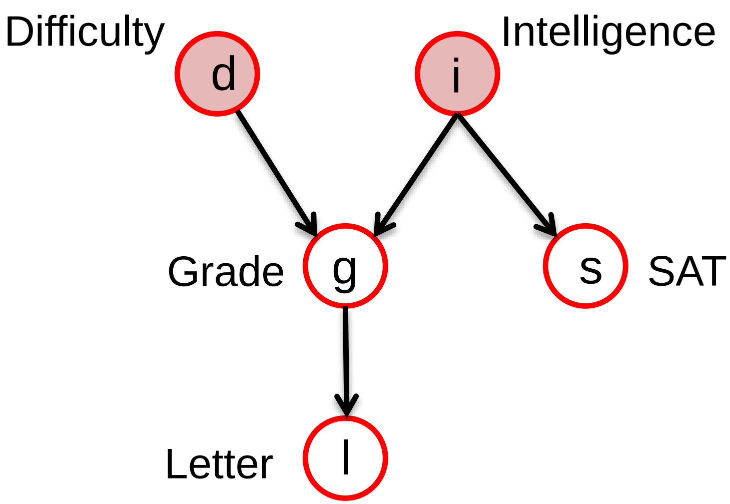
i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's



Path from d to i is blocked. → d,i are independent and p(d,i)=p(d)p(i)

What is the probability of a good letter given that the student is normally intelligent and the exam is easy: p(l=1|i=0,d=0)

$$p(d,i,g,s,l) = p(g,s,l \mid d,i)\,p(d,i)$$
$$= p(d)\,p(i)\,p(g \mid i,d)\,p(s \mid i)\,p(l \mid g)$$

An Example: Now with real p's

What is the probability of a good letter given that the student is normally intelligent and the exam is easy: p(l=1|i=0,d=0)

$$p(g,s,l\,|\,d,i) = p(g\,|\,i,d)\,p(s\,|\,i)\,p(l\,|\,g)$$

Marginalizing the conditional probability

$$p(l=1\,|\,i,d) = \sum_{g,s} p(g\,|\,i,d)\,p(s\,|\,i)\,p(l\,|\,g)$$

Sum over all 2*3 = 6 combinations

$$p(l=1\,|\,i=0,d=0) = 0.513$$

d=0	d=1
0.6	0.4

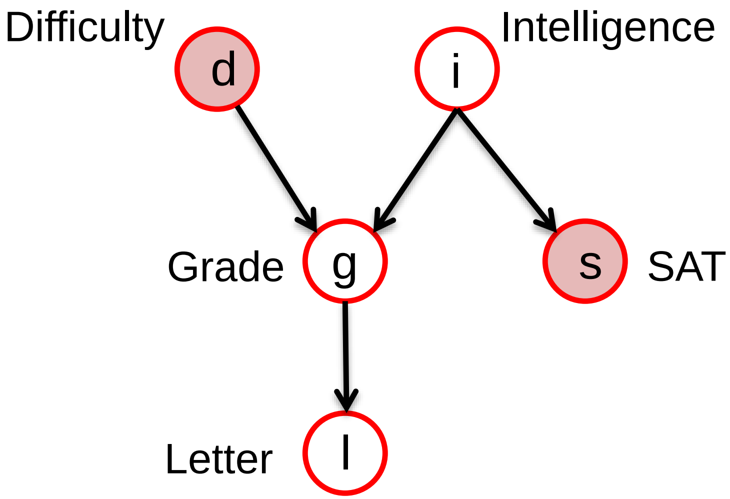
i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's



What the probability of a good letter given that you have a high Sat score and the exam is easy

p(l=1|s=1,d=0)

$$p(d,i,g,s,l) = p(g,s,l \mid d,s)\,p(d,s)$$

$$p(d,i,g,s,l) = p(g,s,l \mid d,s)\,p(d)\,p(s)$$

Path from d to s is blocked by g. → d,s are independent and p(d,s)=p(d)p(s)

An Example: Now with real p's

What the probability of a good letter given that you have a high Sat score and the exam is easy

p(l=1|s=1,d=0)

$$p(d,i,g,s,l) = p(g,i,l \mid d,s)\,p(d)\,p(s)$$

$$p(i,g,l|d,s) = \frac{p(d,i,g,s,l)}{p(d)p(s)}$$

keep in mind:
d,s are inpendent

Joint probability given by the graphical model:

$$p(d,i,g,s,l) = p(d)p(i)p(g|d,i)p(s|i)p(l|g)$$

$$p(i,g,l|d,s) = \frac{\cancel{p(d)}p(i)p(g|d,i)p(s|i)p(l|g)}{\cancel{p(d)}p(s)}$$

$$p(i,g,l|d,s) = \frac{p(i)p(g|d,i)p(s|i)p(l|g)}{\boxed{p(s)}}$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

We need to know p(s)

We have *p(s|i)*, to find *p(s)* we need to marginalize *p(s,i)*.

To get p(s,i) we use the product rule

$$p(s) = \sum p(s,i) = \sum p(s|i)p(i)$$

$$p(s=1) =$$
$$p(s=1|i=0)p(i=0) +$$
$$p(s=1|i=1)p(i=1)$$

$$p(s=1) =$$
$$0.05*0.7 + 0.8*0.3 = 0.275$$

$$p(s=0) = 1 - p(s=1) = 0.725$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

p(s|i)

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

What the probability of a good letter given that you have a high Sat score and the exam is easy

p(l=1|s=1,d=0)

$$p(i,g,l|d,s) = \frac{p(i)p(g|i,d)p(s|i)p(l|g)}{p(s)}$$

Marginalizing the conditional probability

$$p(l=1|s,d) = \sum_{i,g} \frac{p(i)p(g|i,d)p(s|i)p(l|g)}{p(s)}$$

$$= \frac{1}{p(s)} \sum_{i,g} p(i)p(g|i,d)p(s|i)p(l|g)$$

Sum over all 2*3 = 6 combinations

p(l=1|s=1,d=0)

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

An Example: Now with real p's

What the probability of a good letter given that you have a high Sat score and the exam is easy

$$p(l=1|s,d)=\sum_{i,g}\frac{p(i)p(g|i,d)p(s|i)p(l|g)}{p(s)}$$

$$p(l=1|s=0,d=0)=\frac{1}{p(s)}*$$

$$\begin{aligned}(0.7*0.3*0.95*0.9+\\ 0.7*0.4*0.95*0.6+\\ 0.7*0.3*0.95*0.01+\\ 0.7*0.3*0.05*0.9+\\ 0.7*0.4*0.05*0.6+\\ 0.7*0.3*0.05*0.01)\end{aligned}$$

Sum over all 2*3 = 6 combinations

$$p(l=1|s=0,d=0)=0.4953$$

d=0	d=1
0.6	0.4

i=0	i=1
0.7	0.3

	g=1	g=2	g=3
d=0,i=0	0.3	0.4	0.3
d=1,i=0	0.05	0.25	0.7
d=0,i=1	0.9	0.08	0.02
d=1,i=1	0.5	0.3	0.2

	s=0	s=1
i=0	0.95	0.05
i=1	0.2	0.8

	l=0	l=1
g=1	0.1	0.9
g=2	0.4	0.6
g=3	0.99	0.01

Neuroinformatics Lecture

Topics:

Probability Distributions

Probability Distributions

We defined probability as: $p(X = x_i)$

(probability that the random variable X takes the value x_i)

For some cases, we can write down analytically how the probability is dependent on x_i. Then we can define a **‚distribution'** analytically.

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

Here: 2 binary variables → c can be 0, 1, or 2

c=0	**: (0,0)**	**p(c=0)**	**: p(0,0)**
c=1	**: (0,1) ; (1,0)**	**p(c=1)**	**: p(0,1)+p(1,0)**
c=2	**: (1,1)**	**p(c=2)**	**: p(1,1)**

Probability Distributions

We defined probability as:

$$p(X = x_i)$$

(probability that the random variable X takes the value x_i)

For some cases, we can write down analytically how the probability is dependent on x_i. Then we can define a ‚**distribution**' analytically.

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0) : p(0,0)
p(c=1) : p(0,1)+p(1,0)
p(c=2) : p(1,1)

p(0,0) = p(0)p(0) **(the two flips are independent)**
p(0,1) = p(0)p(1)
p(1,0) = p(0,1)
p(1,1) = p(1)p(1)

Probability Distributions

We defined probability as: $p(X = x_i)$

(probability that the random variable X takes the value x_i)

For some cases, we can write down analytically how the probability is dependent on x_i. Then we can define a **‚distribution‘** analytically.

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What‘s the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0) : p(0,0)
p(c=1) : p(0,1)+p(1,0)
p(c=2) : p(1,1)

p(0,0) = p(0)p(0) = 0.5 * 0.5 = 0.25
p(0,1) = p(0)p(0) = 0.5 * 0.5 = 0.25
p(0,1) = p(1,0) = 0.5 * 0.5 = 0.25
p(1,1) = p(1)p(1) = 0.5 * 0.5 = 0.25

Probability Distributions

We defined probability as: $$p(X = x_i)$$

(probability that the random variable X takes the value x_i)

For some cases, we can write down analytically how the probability is dependent on x_i. Then we can define a **‚distribution'** analytically.

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0)	**: p(0,0)**	**= 0.25**
p(c=1)	**: p(0,1)+p(1,0)**	**= 0.25+0.25 = 0.5**
p(c=2)	**: p(1,1)**	**= 0.25**

p(0,0) = p(0)p(0)	**= 0.5 * 0.5 = 0.25**
p(0,1) = p(0)p(0)	**= 0.5 * 0.5 = 0.25**
p(0,1) = p(1,0)	**= 0.5 * 0.5 = 0.25**
p(1,1) = p(1)p(1)	**= 0.5 * 0.5 = 0.25**

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0)	**: p(0,0)**	**= 0.25**
p(c=1)	**: p(0,1)+p(1,0)**	**= 0.25+0.25 = 0.5**
p(c=2)	**: p(1,1)**	**= 0.25**

An anlytical expression for that probability: the binominal distribution:

Probability to get c heads for N=2 flips given that the probability of a single head is μ

$$p(c|\mu, N) = J \cdot \mu^c (1-\mu)^{N-c}$$

Analytical definition of the ‚**distribution**'

c=0	: (0,0)	J= 1
c=1	: (0,1) ; (1,0)	J= 2
c=2	: (1,1)	J= 1

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0) : p(0,0) = 0.25
p(c=1) : p(0,1)+p(1,0) = 0.25+0.25 = 0.5
p(c=2) : p(1,1) = 0.25

An anlytical expression for that probability: the binominal distribution:
Probability to get c heads for N=2 flips given that the probability of a single head is μ

$$p(c|\mu, N) = J \cdot \mu^{c}(1-\mu)^{N-c}$$

$$p(c=0|\mu=0.5, N=2) = 1 \cdot 0.5^{0}(1-0.5)^{2} = 0.25$$

c=0	: (0,0)	J= 1
c=1	: (0,1) ; (1,0)	J= 2
c=2	: (1,1)	J= 1

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0)	**: p(0,0)**	**= 0.25**
p(c=1)	**: p(0,1)+p(1,0)**	**= 0.25+0.25 = 0.5**
p(c=2)	**: p(1,1)**	**= 0.25**

An anlytical expression for that probability: the binominal distribution:

Probability to get c heads for N=2 flips given that the probability of a single head is μ

$$p(c|\mu, N) \quad = J \cdot \mu^c (1-\mu)^{N-c}$$

$$p(c=1|\mu=0.5, N=2) \quad = 2 \cdot 0.5^1 (1-0.5)^1 = 2 * 0.25$$

c=0	: (0,0)	J= 1
c=1	: (0,1) ; (1,0)	J= 2
c=2	: (1,1)	J= 1

For example: So far we computed the probability 'by hand':

Suppose we flip a coin and we assume it is a fair coin, i.e. p(X= tail) = 0.5

What's the probability that we get c= 0, 1 and 2 tails when we flip the coin twice.

p(c=0) : p(0,0) = 0.25
p(c=1) : p(0,1)+p(1,0) = 0.25+0.25 = 0.5
p(c=2) : p(1,1) = 0.25

An anlytical expression for that probability: the binominal distribution:

Probability to get c heads for N=2 flips given that the probability of a single head is μ

$$p(c|\mu, N) = J \cdot \mu^c (1-\mu)^{N-c}$$

$$p(c=2|\mu=0.5, N=2) = 1 \cdot 0.5^2 (1-0.5)^0 = 0.25$$

c=0	: (0,0)	J= 1
c=1	: (0,1) ; (1,0)	J= 2
c=2	: (1,1)	J= 1

Probability Distributions

We defined probability as: $$p(X = x_i)$$

(probability that the random variable X takes the value x_i)

- **A Discrete Probability Distribution:**

  A distribution is discrete if the random variable X is discrete, such that the range of X is finite or countably infinite.

  e.g. Poisson distribution, Bernoulli distribution, binomial distribution geometric distribution, negative binomial distribution

- **A Continuous Probability Distribution:**

  A distribution of a random variable X is called continuous if the range of X is uncountably infinite.

  e.g. normal, uniform, chi-squared, and others.

Discrete Distributions

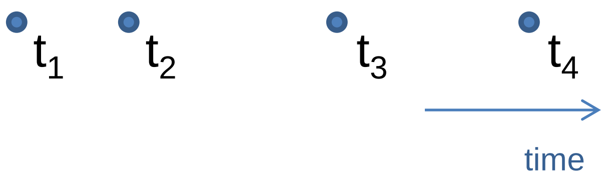
- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

Continuous Distributions

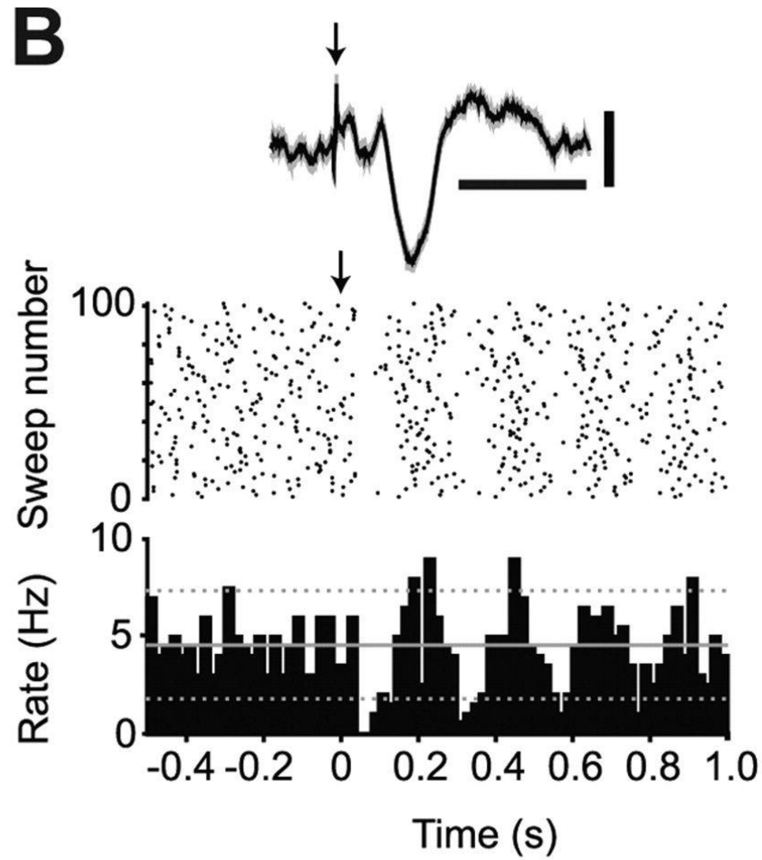
- Normal (Gauss)
- Gamma Distribution
- chi^2 Distribution

Suppose we observe a spiking neuron

The neuron fires at: $t_1, t_2, \ldots, t_N$



We can represent this time by a binary sequence when we bin time in small slices



Individual and population responses of dopaminergic substantia nigra neurons to an aversive electrical stimulus. ***B***, PSTH and raster plot showing the inhibition of an individual dopaminergic neuron in response to stimulus delivery.

c.f. *The Journal of Neuroscience, 4 March 2009, 29(9): 2915-2925; doi: 10.1523/JNEUROSCI.4423-08.2009*

Binary Variables

We begin by considering a single binary random variable $x \in \{0, 1\}$.

For example, x might describe the outcome of flipping a coin, with $x = 1$ representing 'heads', and $x = 0$ representing 'tails'. If the coin is fair, the probability is 0.5 for tails and heads.

In general, we can define a probability of tails and heads given a parameter that we call μ

$$p(x = 1|\mu) = \mu$$

Now we define a distribution (Bernoulli Distribution) that defines the probability of x given the distribution parameter μ

$$p(x|\mu) = \mu^x (1 - \mu)^{1-x}$$

Bernoulli Distribution (one function that describes the probability for all x)

Bernoulli Distribution that defines the probability of x, i.e. x=1 or x=0, given μ

$$p(x|\mu) = \mu^{x}(1-\mu)^{1-x}$$

$$p(x=1|\mu) = \mu^{1}(1-\mu)^{1-1} = \mu$$

$$p(x=0|\mu) = \mu^{0}(1-\mu)^{1-0} = 1-\mu$$

with the probability of a heads defined by the parameter μ

$$p(x=1|\mu) = \mu$$

The distribution is **ONE function** that analytically describes the probability for **all potential outcomes**.

Summary: Bernoulli Distribution

Distribution: $p(x|\mu) = \mu^x(1-\mu)^{1-x}$

Expected value: $E[x] = \mu$

Variance: $var[x] = \mu(1-\mu)$

Y = binopdf(x,n,p)

For n=1, computes the Bernoulli pdf at each of the values in X using the corresponding probability P. The values in P must lie on the interval [0, 1].

Discrete Distributions

- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

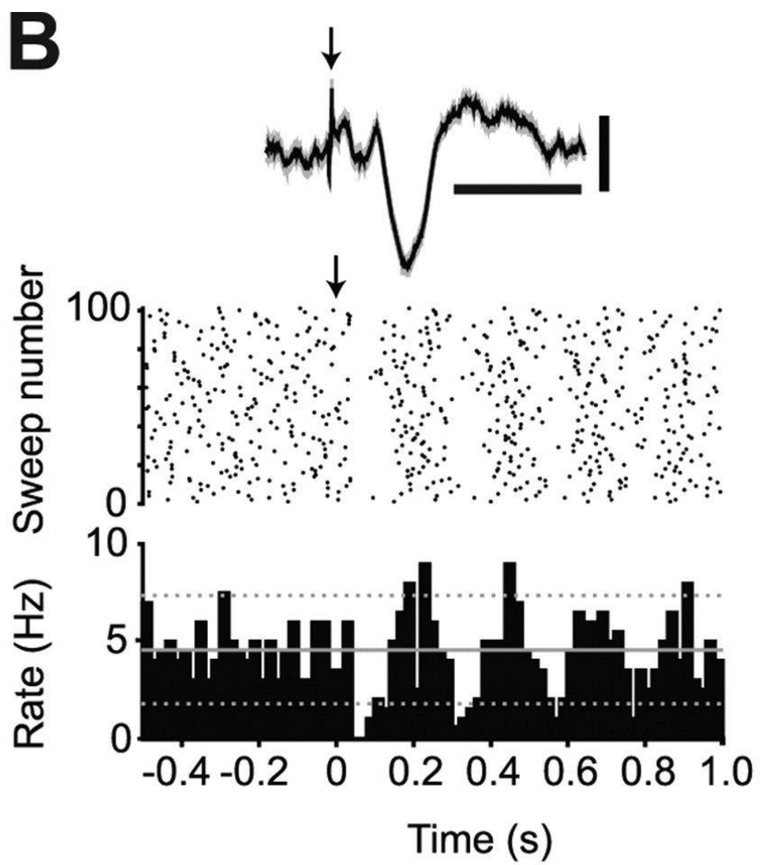
Continuous Distributions

- Normal (Gauss)
- Gamma Distribution
- chi^2 Distribution

Suppose we observe a spiking neuron in many trials.

We can represent every trial with an binary sequence.

T1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
T2	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0
T3	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
T4	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0



Individual and population responses of dopaminergic substantia nigra neurons to an aversive electrical stimulus. ***B***, PSTH and raster plot showing the inhibition of an individual dopaminergic neuron in response to stimulus delivery.

c.f. *The Journal of Neuroscience, 4 March 2009, 29(9): 2915-2925; doi: 10.1523/JNEUROSCI.4423-08.2009*

The binomial distribution is a natural extension of the Bernoulli distribution.

Suppose that we observe a binary random variable X for N times.

Each outcome is Bernoulli distributed: $p(x|\mu) = \mu^{x}(1-\mu)^{1-x}$

The outcome for each random variable is independent of the others. Moreover, the probability is the same, i.e. *i.i.d.* (independent and identically distributed).

i.i.d

1. What does i.i.d mean ?

i.i.d: means the random variables at hand are ***independent*** and ***identically distributed***

2. What does i.i.d. of the random variables a,b,c mean for a factorisation of a joint probability p(a,b,c)

Independence of a,b,c implies : p(a,b,c) = p(a)p(b)p(c)
identically distributed implies: all three random variables a,b,c come form the same distribution. For example from a Bernoulli distribution. And that means that all there distributions are parameterized the same way. In case of the Bernoulli distribution they have the same μ.

$$p(a,b,c) = \overbrace{p(a)\,p(b)\,p(c)}^{\textit{Independence}} = \prod_{i=1}^{3} \overbrace{p(x_i)}^{\textit{identically distributed}}$$

The binomial distribution is a natural extension of the Bernoulli distribution.

Suppose that we observe a binary random variable X for N times.

Each outcome is Bernoulli distributed: $p(x|\mu) = \mu^x(1-\mu)^{1-x}$

The outcome for each random variable is independent of the others. Moreover, the probability is the same, i.e. *i.i.d.* (independent and identically distributed).

What is the probability that we have k heads for N flips ?

What is the probability that we have k heads for N flips ?

For each flip it holds: $p(x|\mu) = \mu^x(1-\mu)^{1-x}$

1. We define a new random variable $D = \{x_1, x_2, \ldots, x_N\}$

2. We now define the probability of *D that is the joint probability of the set of* x_i

$$p(D|\mu) = p(x_1, x_2, \ldots, x_N|\mu)$$

3. All x_i are *i.i.d*, (independent and identically distributed)

$$p(D|\mu) = \prod_{i=1}^{N} p(x_i|\mu) = \prod_{i=1}^{N} \mu^x(1-\mu)^{1-x}$$

So far: x_i are *i.i.d*, $D = \{x_1, x_2, \dots, x_N\}$, and $p(D|\mu) = \prod_{i=1}^{N} \mu^x (1-\mu)^{1-x}$

Actually we are interested in the probability of k that is p(k|μ,N)

(What is the probability that we have k heads for N flips?)

4. There are different D's that are giving the same sum of heads k.

   i.e. 1,1,1,0 / 1,1,0,1 / 1,0,1,1 / 0,1,1,1 has all three heads

$$p(k|\mu, N) = \sum_{j=1}^{J} p\left(D_j|\mu\right) \;\; with\; D_j = \left\{x_1^j, x_2^j, \dots, x_N^j\right\} \;\; and \;\; k = \sum_{i}^{N} x_i^j$$

T1	1	1	0	1	1	1	0	0	1	0	1	1	0	0	0	0
T2	1	1	1	0	1	1	1	0	0	1	0	0	1	0	0	0
T3	1	1	1	1	0	0	1	1	0	0	1	0	0	1	0	0
T4	1	0	1	1	1	0	1	1	1	1	1	0	0	0	1	0
	k=4	k=3				k=2						k=1				k=0

$$p(1|\mu, 4) = p(0{,}0{,}0{,}1|\mu) + p(0{,}0{,}1{,}0|\mu) + p(0{,}1{,}0{,}0|\mu) + p(1{,}0{,}0{,}0|\mu)$$

So far: x_i are *i.i.d*, $D = \{x_1, x_2, \dots, x_N\}$, and $p(D|\mu) = \prod_{i=1}^{N} \mu^x (1-\mu)^{1-x}$

Actually we are interested in the probability of k that is p(k|μ,N)

(What is the probability that we have k heads for N flips?)

4. There are different D's that are giving the same sum of heads k.

   i.e. 1,1,1,0 / 1,1,0,1 / 1,0,1,1 / 0,1,1,1 has all three heads

$$p(k|\mu, N) = \sum_{j=1}^{J} p(D_j|\mu) \ \mathit{with}\ D_j = \{x_1^j, x_2^j, \dots, x_N^j\} \ \ \mathit{and}\ \ k = \sum_i^N x_i^j$$

3. *Since all x_i i.i.d* $p(D_{j=l}|\mu) = p(D_{j=h}|\mu)$ (geeky way of saying: all p are the same)

$$p(k|\mu, N) = J \cdot p(D_j|\mu)$$

So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ *with* $D_j = \{x_1^j, x_2^j, \ldots, x_N^j\}$ *and* $k = \sum_i^N x_i^j$

3. Next we need to know J (number of D‘s that have the same number of heads)

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

reads N choose k

Binomial coefficient:

Example: 2 binary variables → k can be 0, 1, or 2

k=0	: (0,0)	J= 1
k=1	: (0,1) ; (1,0)	J= 2
k=2	: (1,1)	J= 1

So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ $with\ D_j = \{x_1^j, x_2^j, \ldots, x_N^j\}$ $and\ \ k = \sum_i^N x_i^j$

3. *Next we need to know J (number of D's that have the same number of heads)*

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

reads N choose k

Binomial coefficient:

Example: 3 binary variables → k can be 0, 1, 2, or 3

k=0	: (0,0,0)	J= 1
k=1	: (0,0,1) ; (0,0,1); (1,0,0)	J= 3
k=2	: (0,1,1) ; (1,0,1); (1,1,0)	J= 3
k=3	: (1,1,1)	J= 1

So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ *with* $D_j = \{x_1^j, x_2^j, \ldots, x_N^j\}$ *and* $k = \sum_i^N x_i^j$

3. Next we need to know J (number of D's that have the same number of heads)

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

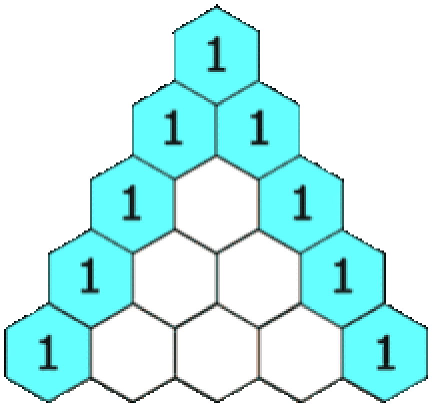
reads N choose k

Pascal's triangle.

1	N=0
1 1	N=1
1 2 1	N=2
1 3 3 1	N=3
1 4 6 4 1	N=4

k=0

k=N



So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ *with* $D_j = \left\{x_1^j, x_2^j, \ldots, x_N^j\right\}$ *and* $k = \sum_i^N x_i^j$

3. *Next we need to know J (number of D's that have the same number of heads)*

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

reads N choose k

Further link: You may still know the binomial coefficient from high school.

$$(x + y)^n = a_0 x^n + a_1 x^{n-1} y + a_2 x^{n-2} y^2 + \ldots + a_{n-1} x y^{n-1} + a_n y^n$$

$$a_j = \binom{n}{j}$$

So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ *with* $D_j = \{x_1^j, x_2^j, \ldots, x_N^j\}$ *and* $k = \sum_i^N x_i^j$

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

reads N choose k

3. *Next we use p(D|μ) from the first steps*

$$p(D|\mu) = \prod_{i=1}^{N} \mu^{x_i}(1-\mu)^{1-x_i}$$

$$p(D|\mu) = \prod_{i=1}^{N} \mu^{x_i} \cdot \prod_{i=1}^{N} (1-\mu)^{1-x_i} \quad \textit{with } \mu^0 = 1 \textit{ and } \mu^1 = \mu$$

$$p(D|\mu) = \prod_{i=1}^{k} \mu \cdot \prod_{i=1}^{N-k} (1-\mu)$$

$$p(D|\mu) = \mu^k (1-\mu)^{N-k}$$

So far: $p(k|\mu, N) = J \cdot p(D_j|\mu)$ $with\ D_j = \{x_1^j, x_2^j, \dots, x_N^j\}$ $and\ k = \sum_i^N x_i^j$

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot N}{1 \cdot 2 \cdot \ldots \cdot k \cdot 1 \cdot 2 \cdot \ldots \cdot (N-k)}$$

reads N choose k

3. *Next we use p(D|μ) from the first steps*

$$p(k|\mu, N) = J\, p(D|\mu)$$
$$with\ p(D|\mu) = \mu^k (1-\mu)^{N-k}$$

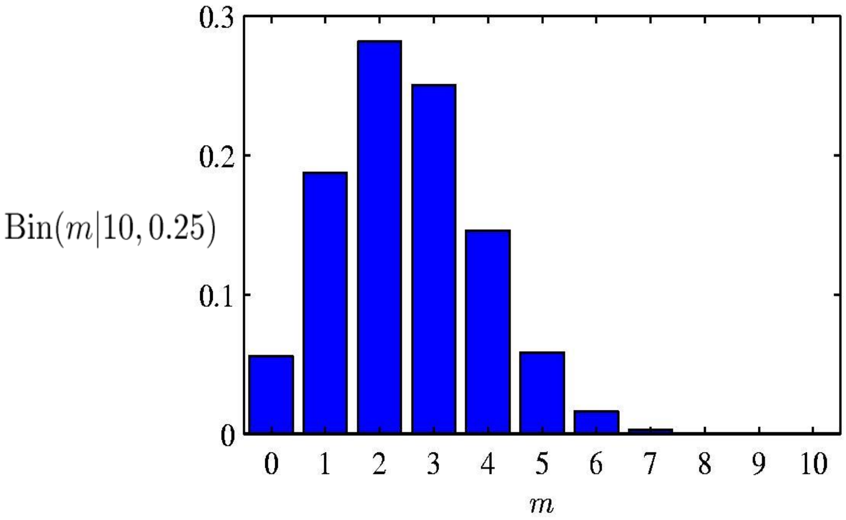
$$p(k|\mu, N) = J\, p(D|\mu) = \binom{N}{k} \cdot \mu^k (1-\mu)^{N-k}$$

That‘s the binominal distribution:

Probability to get k heads for N flips given that the probability of a single flip is μ

$$p(k|\mu, N) = \binom{N}{k} \cdot \mu^k (1-\mu)^{N-k}$$

Binomial Distribution



$$\mathrm{Bin}(m|N,\mu) = \binom{N}{m}\mu^m(1-\mu)^{N-m}$$

$$\mathbb{E}[m] \equiv \sum_{m=0}^{N} m\mathrm{Bin}(m|N,\mu) \quad = \quad N\mu$$

$$\mathrm{var}[m] \equiv \sum_{m=0}^{N} (m-\mathbb{E}[m])^2\,\mathrm{Bin}(m|N,\mu) \quad = \quad N\mu(1-\mu)$$

Y = binopdf(x,n,p)

computes the binomial pdf at each of the values in X using the corresponding parameters in N and P. The parameters in N must be positive integers, and the values in P must lie on the interval [0, 1].

Q

1. What does i.i.d mean ?

2. What does i.i.d. of the random variables a,b,c mean for a factorisation of a joint probability p(a,b,c)

3. What does N choose k mean ? Compute J for N=4 and k=0, k=1, k=2, k=3, k=4

$$J = \binom{N}{k} = \frac{N!}{k!(N-k)!}$$

Timer (8min): Start Stop

1. What does i.i.d mean ?

i.i.d: means the random variables at hand are ***independent*** and ***identically distributed***

2. What does i.i.d. of the random variables a,b,c mean for a factorisation of a joint probability p(a,b,c)

Independence of a,b,c implies : p(a,b,c) = p(a)p(b)p(c)
identically distributed implies: all three random variables a,b,c come form the same distribution. For example from a Bernoulli distribution. That means that all there distribution a parameterized the same way. In case of the Bernoulli distribution they have the same μ.

$$p(a,b,c) = \overbrace{p(a)\,p(b)\,p(c)}^{\textit{Independence}} = \prod_{i=1}^{3} \overbrace{p(x_i \mid \mu)}^{\textit{identically distributed}}$$

Conditional Independence

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}$$

$$\binom{4}{0} = \frac{1\cdot 2\cdot 3\cdot 4}{0!(1\cdot 2\cdot 3\cdot 4)} = 1$$

$$\binom{4}{1} = \frac{1\cdot 2\cdot 3\cdot 4}{1!(1\cdot 2\cdot 3)} = \frac{4}{1!} = 4$$

$$\binom{4}{2} = \frac{1\cdot 2\cdot 3\cdot 4}{1\cdot 2(1\cdot 2)} = \frac{3\cdot 4}{1\cdot 2} = 6$$

$$\binom{4}{3} = \frac{1\cdot 2\cdot 3\cdot 4}{1\cdot 2\cdot 3(1)} = \frac{4}{1} = 4$$

$$\binom{4}{4} = \frac{1\cdot 2\cdot 3\cdot 4}{1\cdot 2\cdot 3\cdot 4(0!)} = \frac{1}{1} = 1$$