Neuroinformatics Lecture (L7)

Prof. Dr. Gordon Pipa

Institute of Cognitive Science University of Osnabrück

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.

Summary: Bernoulli Distribution

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

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

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

Discrete Distributions

- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

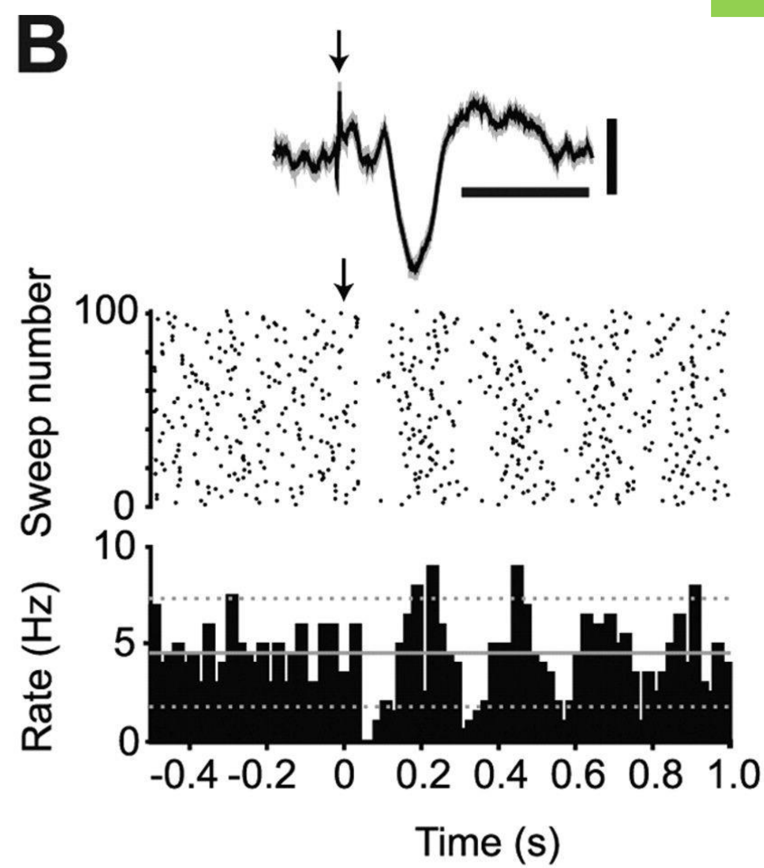
Continuous Distributions

- 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).

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

R

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 $D = \{x_1, x_2, \dots, x_N\}$

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

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

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

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

	k=4
T1	1
T2	1
T3	1
T4	1

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

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

	k=1			
T1	1	0	0	0
T2	0	1	0	0
T3	0	0	1	0
T4	0	0	0	1

	k=0
T1	0
T2	0
T3	0
T4	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*, $D = \{x_1, x_2, \ldots, x_N\}$, 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

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

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

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* $D_j = \{x_1^j, x_2^j, \dots, x_N^j\}$ *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)*

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

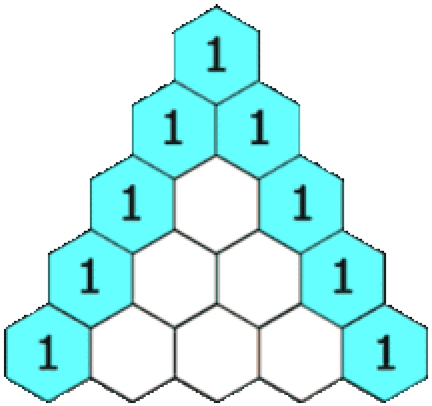
reads N choose k

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

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

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

reads N choose k

3. *Next we use p(D|μ) from the first steps* $p(D|\mu) = \prod_{i=1}^{N} \mu^{x_i}(1-\mu)^{1-x_i}$

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

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

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

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

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

reads N choose k

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

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

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

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

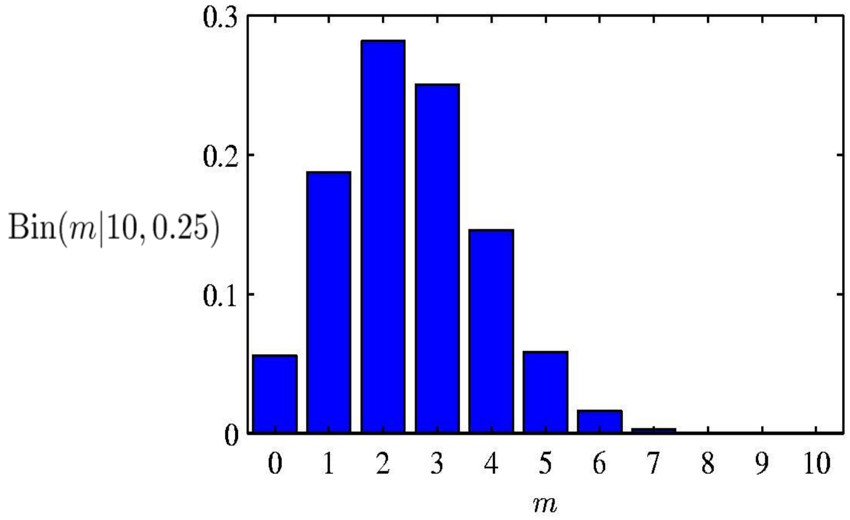
That‘s the binominal distribution:

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.

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

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

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

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

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

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

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

Discrete Distributions

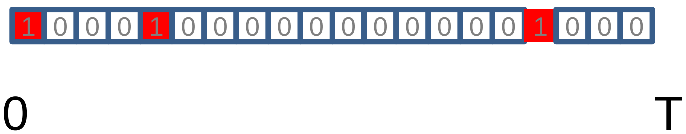
- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

Continuous Distributions

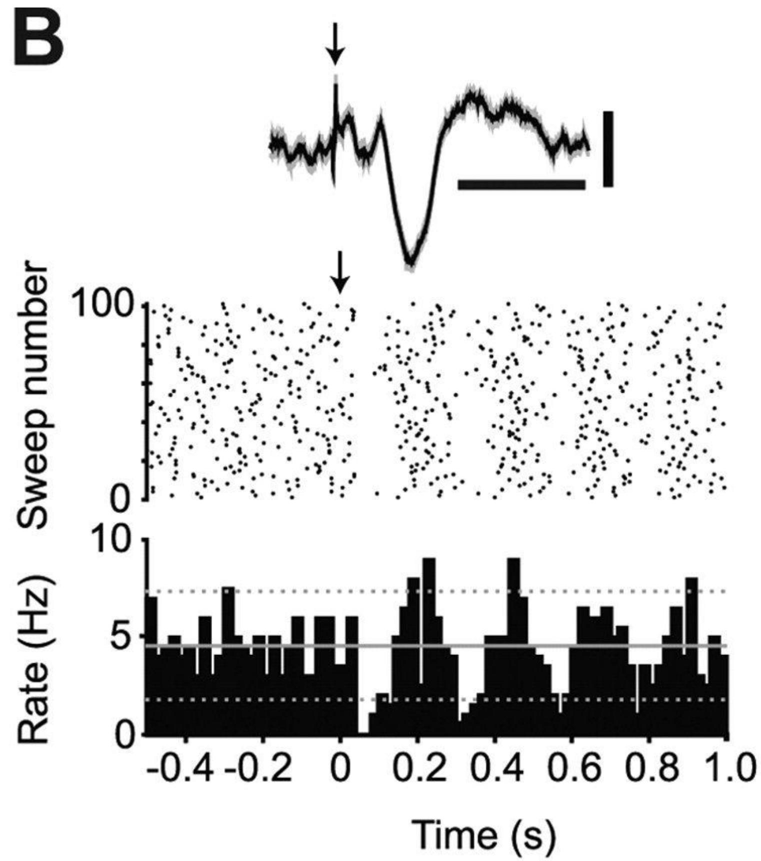
- Normal (Gauss)
- Gamma Distribution
- chi^2 Distribution

Suppose we observe a spiking neuron in many trials.

We are now interested in the probability to have a certain number of spikes in a fixed interval T.

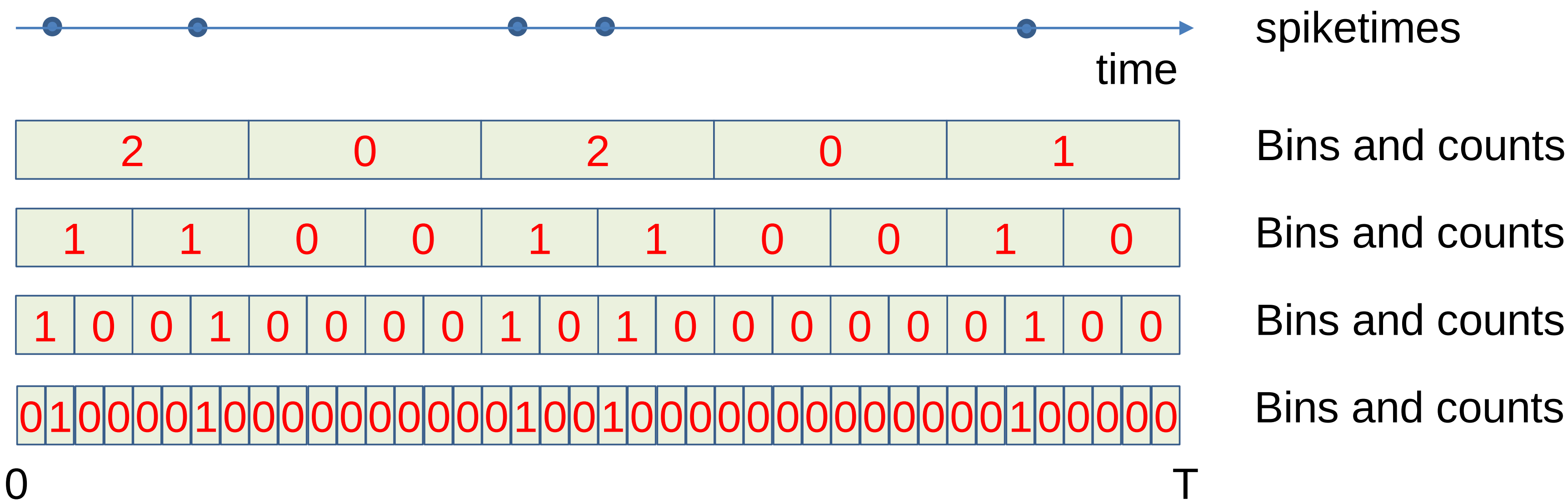


If we make the bins very small, we can approximate the binomial distribution (k events in N bins) with a Poisson distribution.



Individual and population responses of dopamin-ergic 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*



If we make the bins smaller, we get more bins with 0 counts, while the total number of spikes (k) stays the same → Probability per bin μ drops fast

$$\mu = \frac{number\ of\ spikes}{number\ of\ bins} \qquad p(k|\mu, N) = \binom{N}{k} \cdot \mu^k (1-\mu)^{N-k}$$

However, E(k) = N*μ = constant (N = number of bins)

We begin by considering the standard binomial distribution

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

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

We rewrite the binomial as the product of four terms (using $\lambda = N\mu$):

We start with the binomial coefficient:

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

N>K

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

$$= \frac{(N-k+1) \cdot \ldots \cdot (N-1) \;\; \cdot N}{k!}$$

So far: standard binomial distribution

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

We rewrite the dependence on μ using $\lambda = N\mu$:

$$\mu^k (1-\mu)^{\mathrm{N-k}} = \frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^{N-k} = \frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^{N}\left(1-\frac{\lambda}{N}\right)^{-k}$$

Combining both:

$$p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1)\ \cdot N}{k!} \cdot \frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^{N}\left(1-\frac{\lambda}{N}\right)^{-k}$$

So far: $p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1)\ \cdot N}{k!} \cdot \frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^N \left(1-\frac{\lambda}{N}\right)^{-k}$

Now we assume that we use this model to describe a spike train using very small temporal bins to describe the spike train via a binary process.

1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

0 T

We now reduce the bin size to a arbitrary small value which is reducing the probability per bin (limes of p going to zero).

At the same time the number of bins must grow if we want to cover the same temporal range between 0 and T (limes of N going to infinity).

Moreover, we know that, irrespective of the bin size and the number of bins, the expected value of events in the time interval from 0 to T stays the same.

We achieve this by making N large, μ small and keeping the product of μ and N constant at the same time.

So far: $$p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1) \cdot N}{\boxed{k!}} \cdot \frac{\boxed{\lambda^k}}{N^k}\left(1-\frac{\lambda}{N}\right)^N \left(1-\frac{\lambda}{N}\right)^{-k}$$

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ np=const=\lambda}} \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!}$$

Independent of N and p, given λ is constant

So far: $$p(k|\mu,N)=\frac{(N-k+1)\cdot\ldots\cdot(N-1)\ \cdot N}{k!}\cdot\frac{\lambda^k}{N^k}\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-k}$$

$$\lim_{\substack{N\to\infty\\ p\to 0\\ np=\text{const}=\lambda}}\frac{\lambda^k}{k!}=\frac{\lambda^k}{k!}$$

Independent of N and p given λ is constant

$$\lim_{N\to\infty}\frac{N(N-1)\cdot\ldots\cdot(N-k+1)}{N^k}=1$$

If N grows, it becomes ~ N to the power of k

So far: $$p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1)\ \cdot N}{k!} \cdot \frac{\lambda^k}{N^k} \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^{-k}$$

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ np=\text{const}=\lambda}} \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!}$$

Independent of N and p given λ is constant

$$\lim_{N\to\infty} \frac{N(N-1)\cdot \ldots \cdot (N-k+1)}{N^k} = 1$$

If N grows, it becomes ~ N to the power of k

$$\lim_{\substack{N\to\infty \\ Np=\text{const}=\lambda}} \left(1 - \frac{\lambda}{N}\right)^N = e^{-\lambda}$$

known limit (see textbook)

So far: $$p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1) \cdot N}{k!} \cdot \frac{\lambda^k}{N^k} \left(1-\frac{\lambda}{N}\right)^N \left(1-\frac{\lambda}{N}\right)^{-k}$$

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ np=\text{const}=\lambda}} \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!}$$

Independent of N and p, given λ is constant

$$\lim_{N\to\infty} \frac{N(N-1)\cdot\ldots\cdot(N-k+1)}{N^k} = 1$$

If N grows, it becomes ~ N to the power of k

$$\lim_{\substack{N\to\infty \\ np=\text{const}=\lambda}} \left(1-\frac{\lambda}{N}\right)^N = e^{-\lambda}$$

known limit (see textbook)

$$\lim_{\substack{N\to\infty \\ np=\text{const}=\lambda}} \left(1-\frac{\lambda}{N}\right)^{-k} = 1$$

λ over N becomes zero for constant k

So far: $$p(k|\mu, N) = \frac{(N-k+1)\cdot \ldots \cdot (N-1)\ \cdot N}{k!} \cdot \frac{\lambda^k}{N^k} \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^{-k}$$

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ np=\text{const}=\lambda}} p(k\,|\,\mu, N) = \left(\frac{\lambda^k}{k!}\right)\cdot(1)\cdot\left(e^{-\lambda}\right)\cdot(1) = \frac{\lambda^k}{k!}e^{-\lambda}$$

Binary Variables

Poisson Distribution: probability of k events given that λ events are expected

$$p(k|\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$$

With k=0,1,2, …. (discrete)   and λ >0 (parameter is continuous)

E(k)   = λ
var(k)   = λ

Example: p(k) spikes in an interval between 0 and T, if λ spikes are expected

Y = poisspdf(X,lambda)

computes the Poisson pdf at each of the values in X using the corresponding parameters in lambda. The parameters in lambda must all be positive.

Q

1. How many parameters does the Binominal and Poisson distribution have

2. What is/are the basic limit/s that we used for the derivation of the Poisson distribution

$$\lim_{?} \quad p(k \mid \mu, N) = \left(\frac{\lambda^{k}}{k!}\right) \cdot (1) \cdot \left(e^{-\lambda}\right) \cdot (1) = \frac{\lambda^{k}}{k!} e^{-\lambda}$$

3. What the limit of $\lim_{\substack{N \to \infty \\ np = \text{const} = \lambda}} \left(1 - \frac{\lambda}{N}\right)^{-k} = ?$

Timer (5min): Start Stop

1. How many parameters does the Binominal and Poisson distribution have

Poisson Distribution: probability of k events given that λ events are expected

$$p(k|\lambda) = \frac{\lambda^k}{k!}e^{-\lambda}$$

With k=0,1,2, …. and λ >0 (parameter is continuous)

Binominal Distribution: probability of m events with $N\mu$ events expected

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

2. What is/are the basic limit/s that we used for the derivation of the Poisson distribution

$$\lim_{\substack{N\to\infty \\ p\to 0 \\ np=\text{const}=\lambda}} p(k\,|\,\mu, N) = \left(\frac{\lambda^{k}}{k!}\right)\cdot(1)\cdot\left(e^{-\lambda}\right)\cdot(1) = \frac{\lambda^{k}}{k!}e^{-\lambda}$$

3. What the limit of

$$\lim_{\substack{N\to\infty \\ np=\text{const}=\lambda}} \left(1-\frac{\lambda}{N}\right)^{-k} = 1$$

Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

ELSEVIER

Neurocomputing 52–54 (2003) 31–37

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Non-parametric significance estimation of joint-spike events by shuffling and resampling

Gordon Pipa*, Sonja Grün[1]

Department of Neurophysiology, Max-Planck-Institute for Brain Research, Deutschordenstr 46, Frankfurt/Main 60528 Germany

Abstract

The ‘unitary event’ analysis method was designed to analyze multiple parallel spike trains for correlated activity. The null-hypothesis assumes Poissonian spike train statistics, however experimental data may fail to be consistent with this assumption. Here we present a non-parametrical significance test that considers the original spike train structure of experimental data. The significance of coincident events observed in simultaneously recorded spike trains is estimated on the basis of the same spike trains, however recombined to sets of non-corresponding trials (‘trial shuffling’). Resampling from this set provides the distribution for coincidences reflecting the null-hypothesis of independence.

How does the brain compute ?

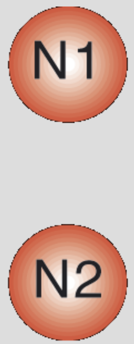
4W

1	hypothesis : single cell code

- information is coded by single neuron

  (e.g. rate of spikes)

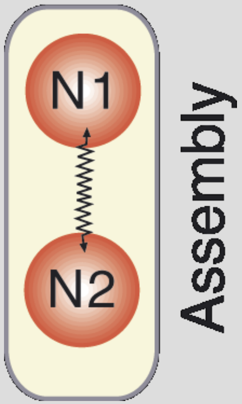
Statistical Hypothesis: **patterns across neurons occur by chance**



2	hypothesis : assembly code

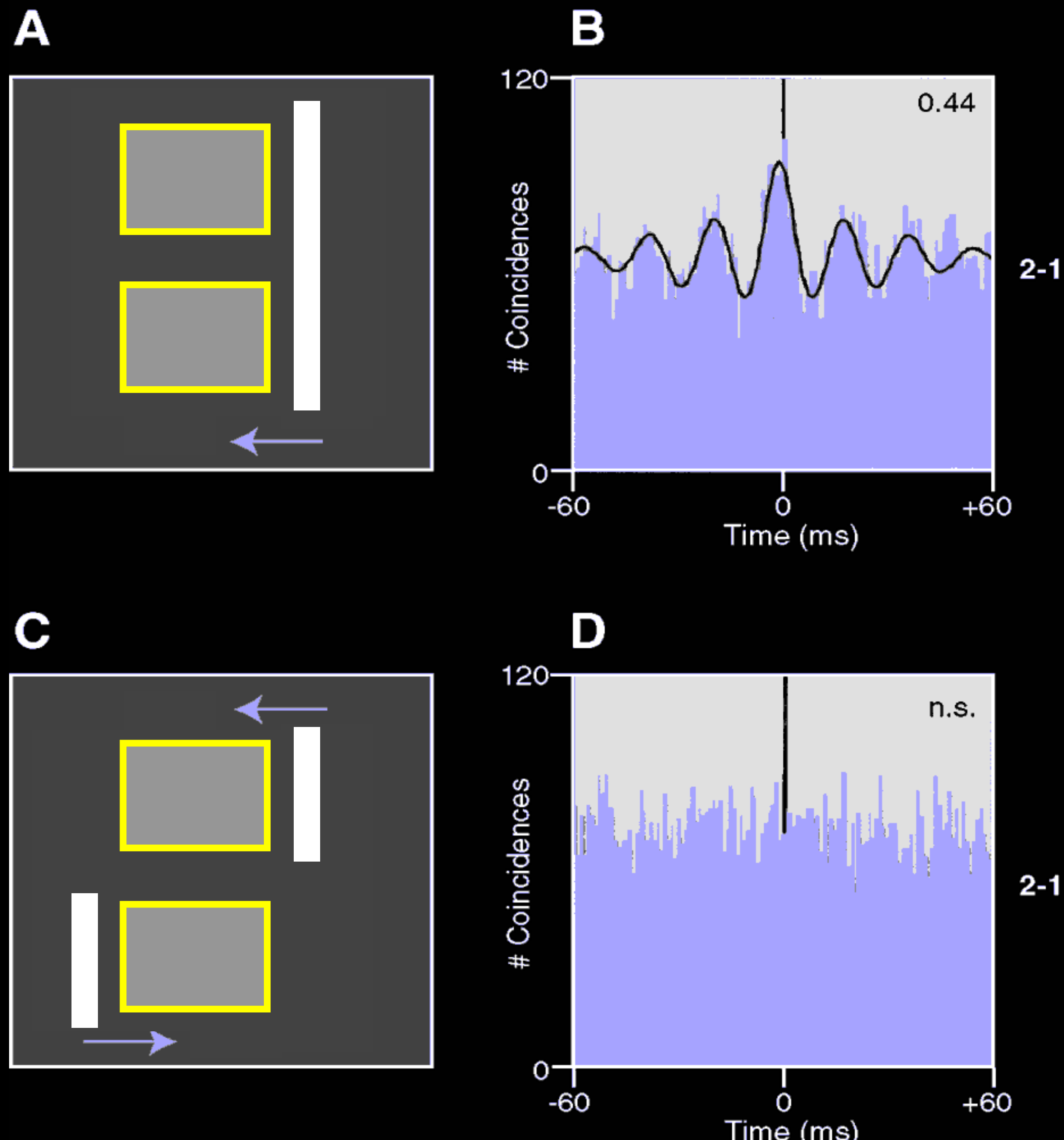
- Information is coded by groups of neurons - assemblies
- Relationship between members of an assembly is expressed by coordinated spiking activity → intrinsic pattern generation

Statistical Hypothesis: **patterns across neurons occur more frequently than expected by chance**



1) Hebb (1949) : 'Organisation of behaviour. A neurophysiological theory' , New York: John Wiley & Sons
2) Singer et al (1997): 'Neuronal assemblies: Necessity, signature and detectability', Trends in Cognitive Science, 252-261

Dynamic relation between neurons



Gray, Singer

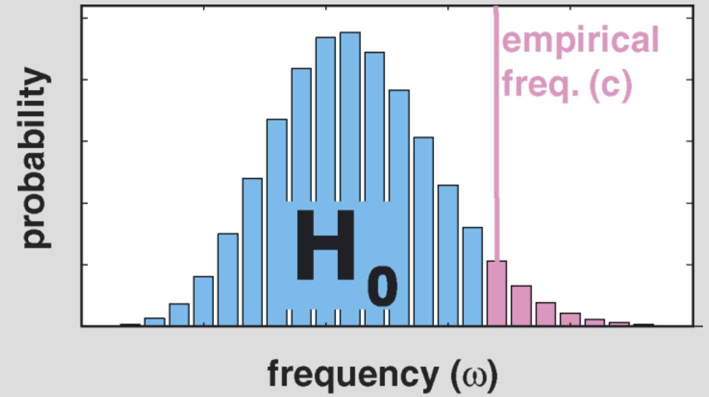
Hypothesis test

Significance Test

H_0 – Hypothesis : pattern occur by chance

H_1 – Hypothesis : pattern occur more frequently than expected by chance

H_0: probability distribution p(ω) required



Unitary Events Method: Parametric Test

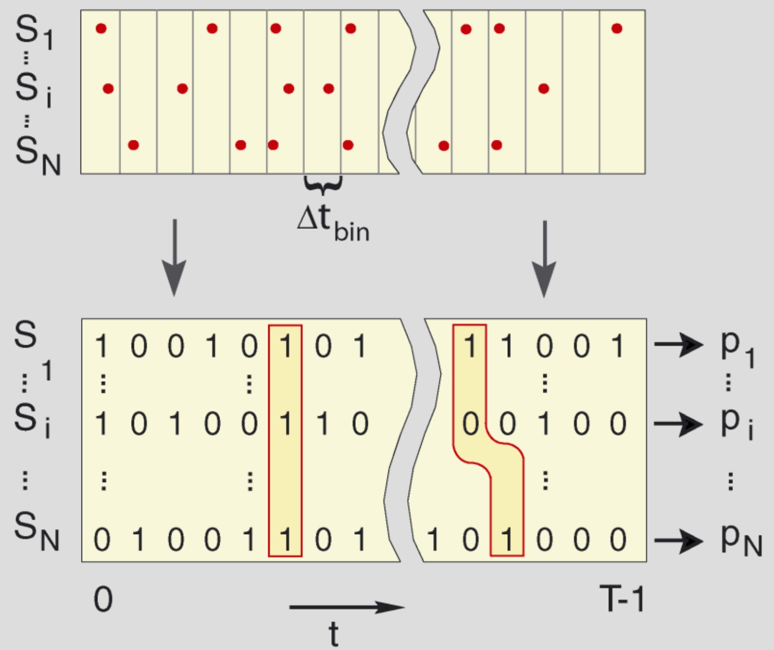
Assumption : Spiketrains → Bernoulli process

$$P = \prod_{i=1}^{N} \varphi_i \quad with \quad \begin{cases} \varphi_i = p_i & \text{spike} \\ \varphi_i = 1 - p_i & \text{no spike} \end{cases}$$

Approximation for P<< 1 :

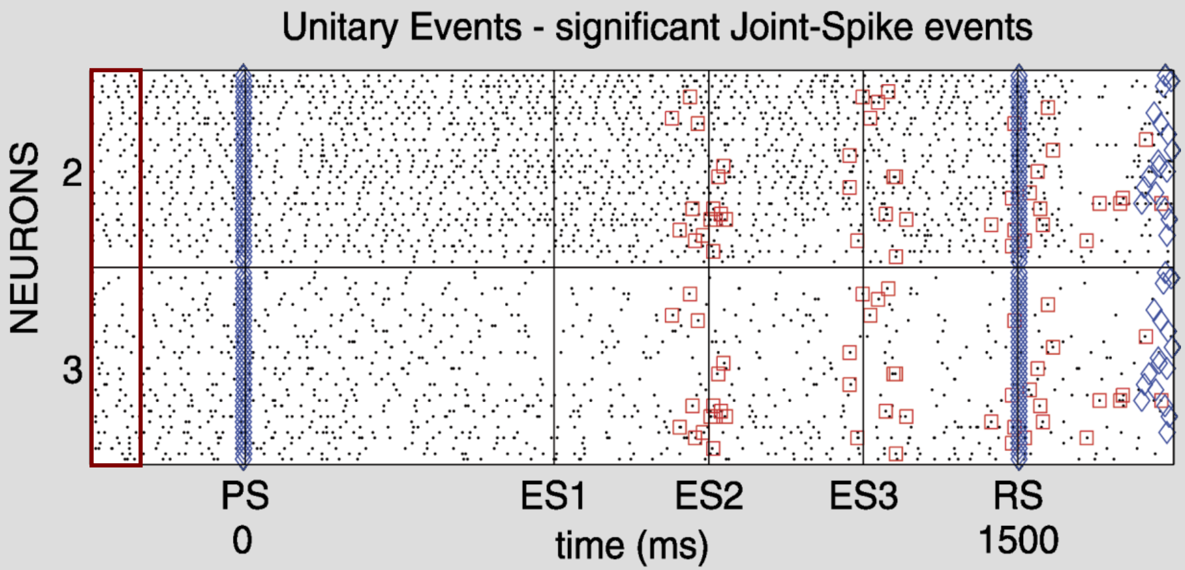
H_0 : Poisson distributed p(ω)

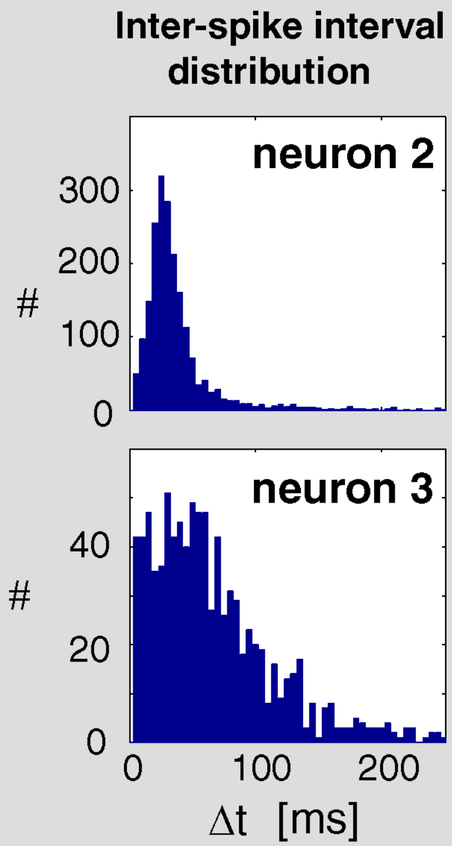
parameterized by : $\langle \omega \rangle = T \cdot P$



Grün et al (2002) : 'Unitary Events in multiple single neuron activity I / II' , Neural Comp. 14, I:43-81; II:82-119

Hypothesis test



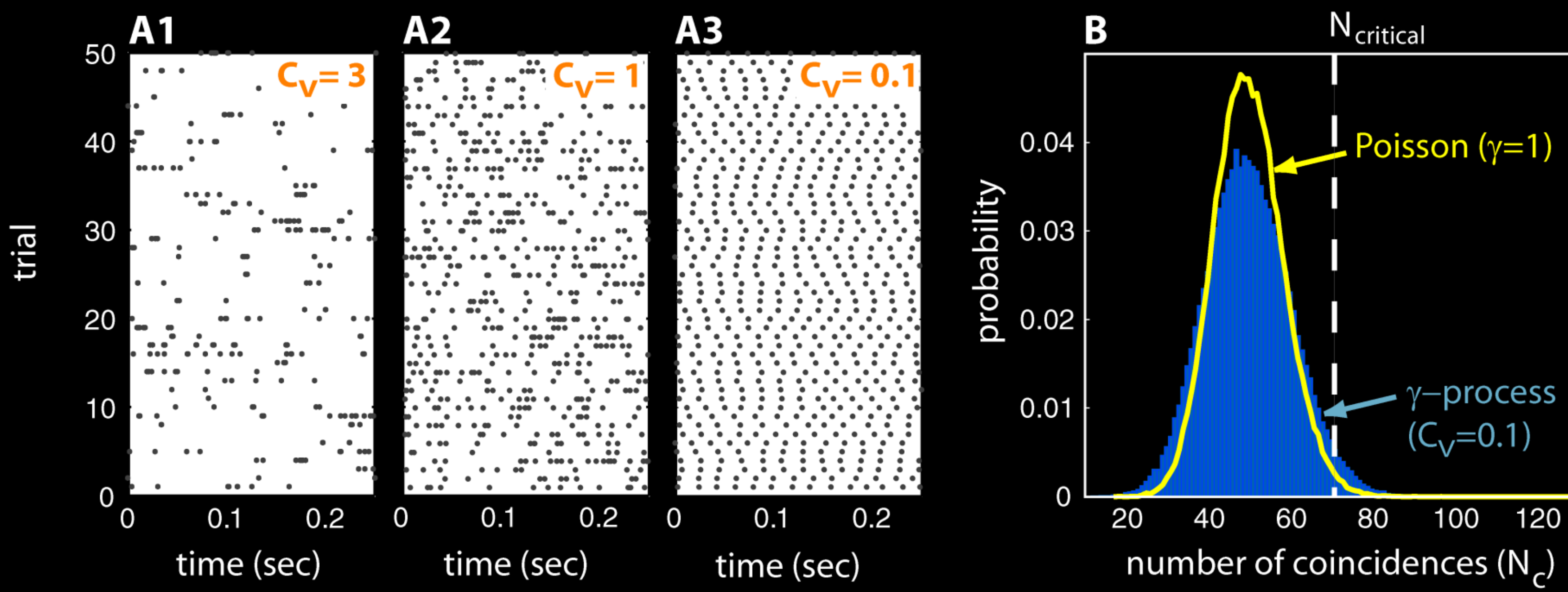


Data : monkey primary motor cortex - 38 trials: delayed-pointing task

Analysis : coincidence window: 5ms - test level: 5%

Riehle, Grün, Diesmann, Aertsen (1997) Science 278: 1950-1953

Impact of Auto-structure on distribution of joint firing (coincidences)



The Dispute: Rate versus Synchrony

4W

Pro temporal coordination

Gray (89): Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex.

"The results demonstrate that local neuronal populations in the visual cortex engage in stimulus-specific synchronous oscillations resulting from an intracortical mechanism. "

Baker (2001): Synchronization in monkey motor cortex during a precision grip task. I. Task-dependent modulation in single-unit synchrony.

"This study provides strong support for assemblies of neurons being synchronized during specific phases of a complex task …"

Castelo-Branco (2002) Neural synchrony correlates with surface segregation rules

„…Thus, dynamic changes in synchronization could encode, in a context-dependent way, relations among simultaneous responses to spatially superimposed contours …"

Contra temporal coordination

Shadlen (98): The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding

"Detailed information about the temporal pattern of synaptic inputs cannot be recovered from the pattern of output spikes, …"

Baker (2000): Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance levels.

„The task dependence of pattern occurrence is explicable as an artifact of the modulation of neural firing rate."

Brody (99): Correlations without synchrony

"Peaks in spike train correlograms are usually taken as indicative of spike timing synchronization between neurons. Strictly speaking, however, a peak merely indicates that the two spike trains were not independent."

How can that happen? Principles of statistical inference

Inference based on a significance test:

~~H_0: Cells fire independently~~

Evidence for H1 ? Yes

H_1: Firing is synchronized

~~H_0: Cells fire independently~~

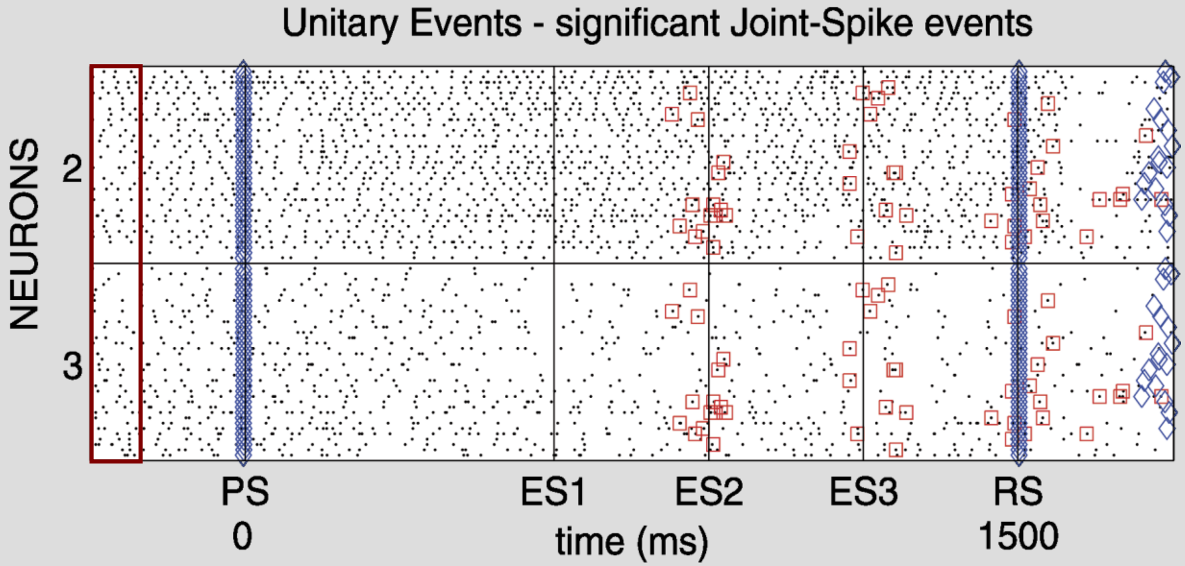
Evidence for H1 ? No

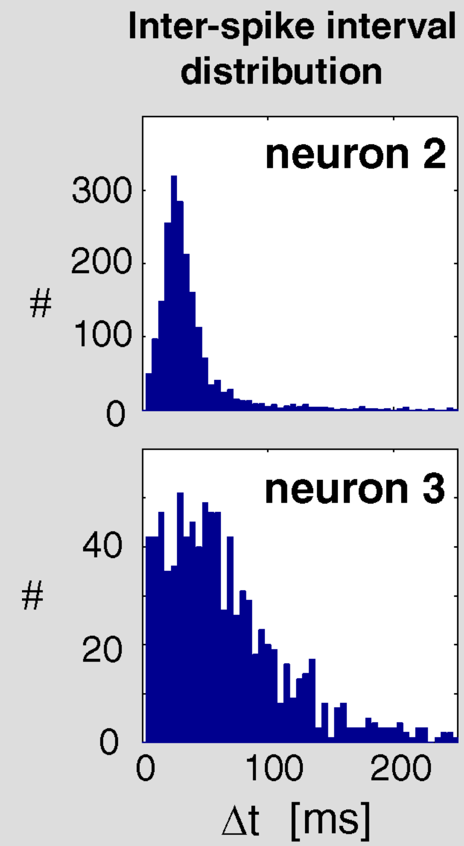
H_1: Firing is synchronized

H_1': Rate co-variation

The significance test must be designed to allow for only one alternative hypothesis.

Hypothesis test





Data : monkey primary motor cortex - 38 trials: delayed-pointing task

Analysis : coincidence window: 5ms - test level: 5%

Riehle, Grün, Diesmann, Aertsen (1997) Science 278: 1950-1953

Non-Parametric Significance test

H_0 – Hypothesis : Deriving H_0 that reflects $p(\omega)$ based on experimental data → considering original spike-train structure	**C S R – Method** **C**ombined **S**huffling and **R**esampling

principle

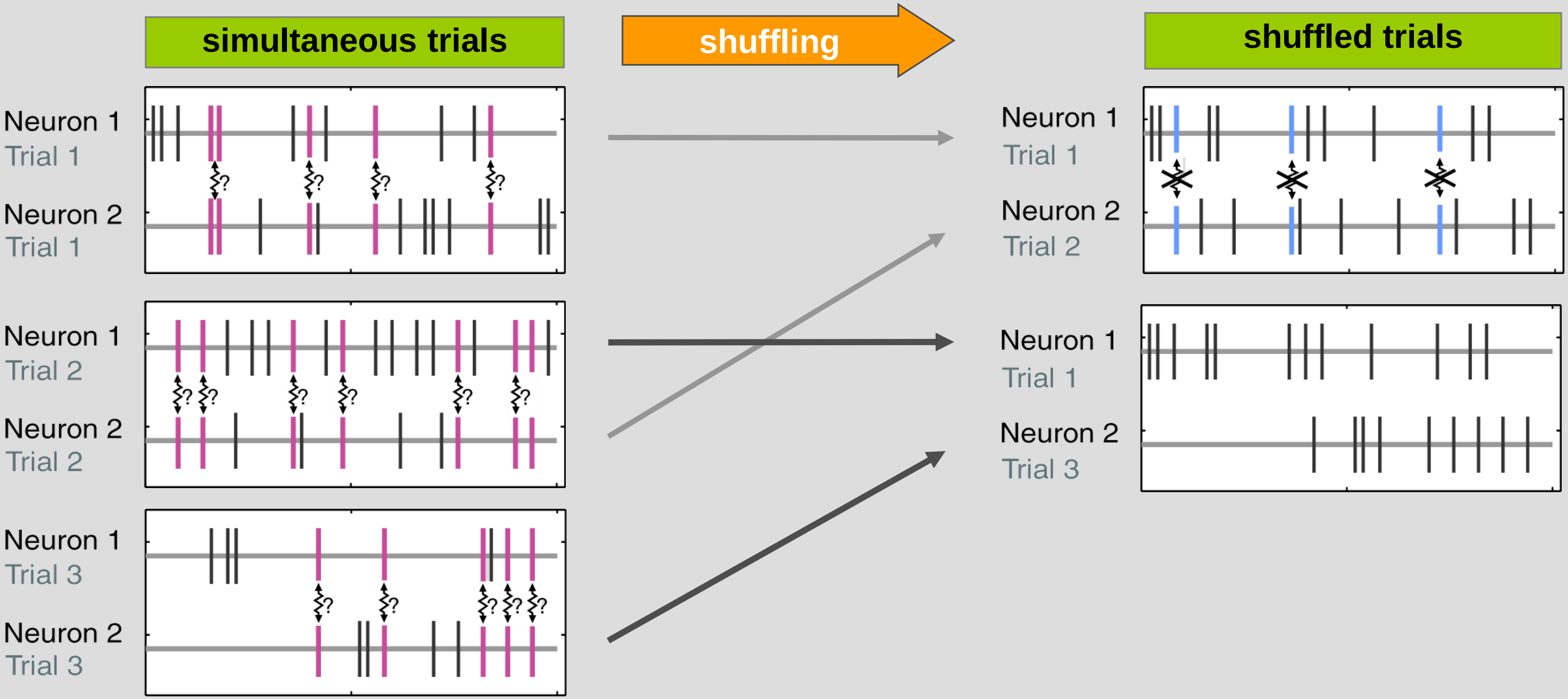
4W

1 **Trial shuffling**	**Intentional destruction of intrinsic pattern**
2 **Resampling**	**Estimation of statistical significance (α) based on shuffled data**

1 Trial shuffling :

Intentional destruction of intrinsic pattern

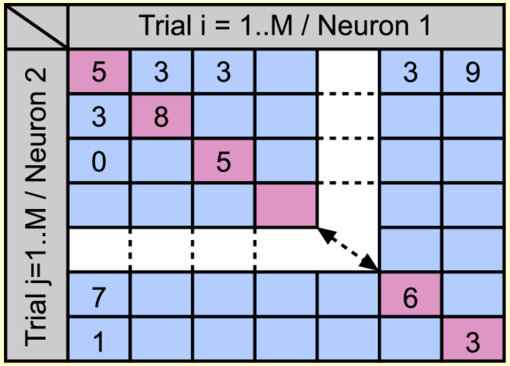
1 Trial shuffling : Destruction of intrinsic pattern



Two Data classes

Ω — **potentially correlated :** (simultaneous)

Ω_0 — **uncorrelated :** (shuffled)

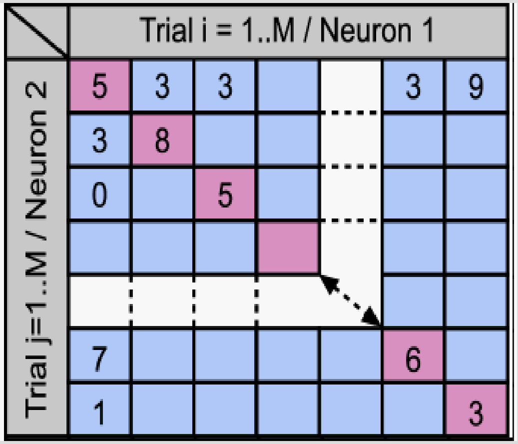


2

Monte Carlo Resampling

Estimate statistical significance based on shuffled trials

Monte-Carlo based significance estimation



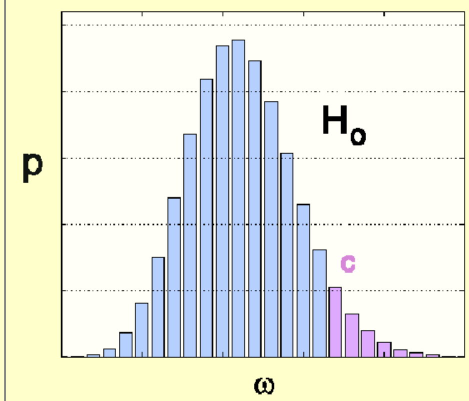
Statistical significance of c based on H_0 Hypothesis

Ω Frequencies of a pattern occurred in M tuples of spike trains

$$c = \sum_{i=1}^{M} \omega_i \quad with \quad \omega_i \in \Omega$$

Ω_0 Frequencies of a pattern occurred in M tuples of shuffled spike trains

$$\omega^* = \sum_{i=1}^{M} \omega_i \quad with \quad \omega_i \in \Omega_0$$



statistical significance :

$$\alpha = p(\omega^* \geq c)$$

Resampling of Ω_0

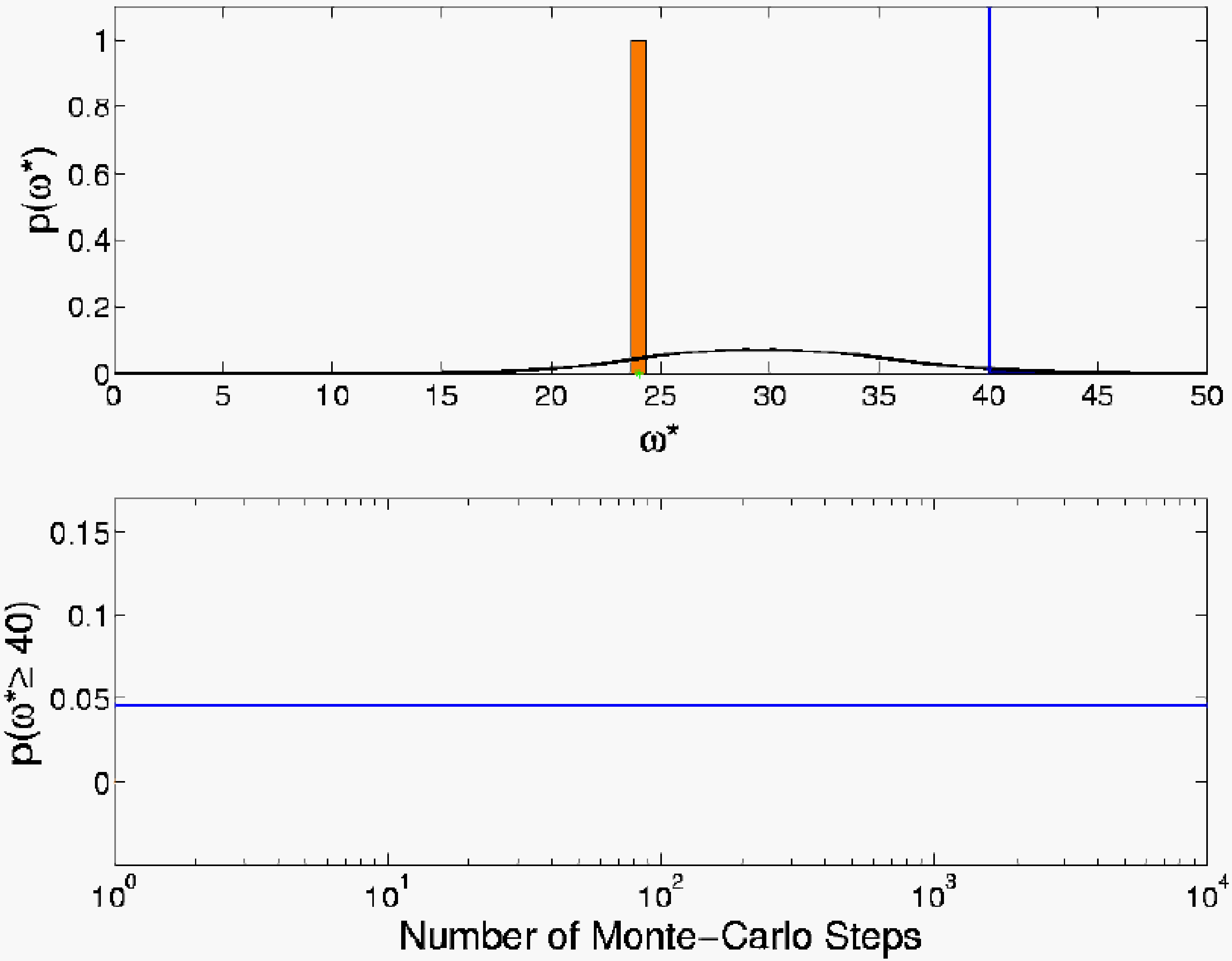
	1	2	3	4	5		M	$\sum$
random-sample 1	9	5	3	7	7		3	ω_1^*
random-sample 2	1	1	1	1	1		1	ω_2^*
random-sample ξ	5	5	9	7	1		1	ω_ξ^*

Monte-Carlo estimation

Estimated significance :
(ξ Bootstrap-samples)

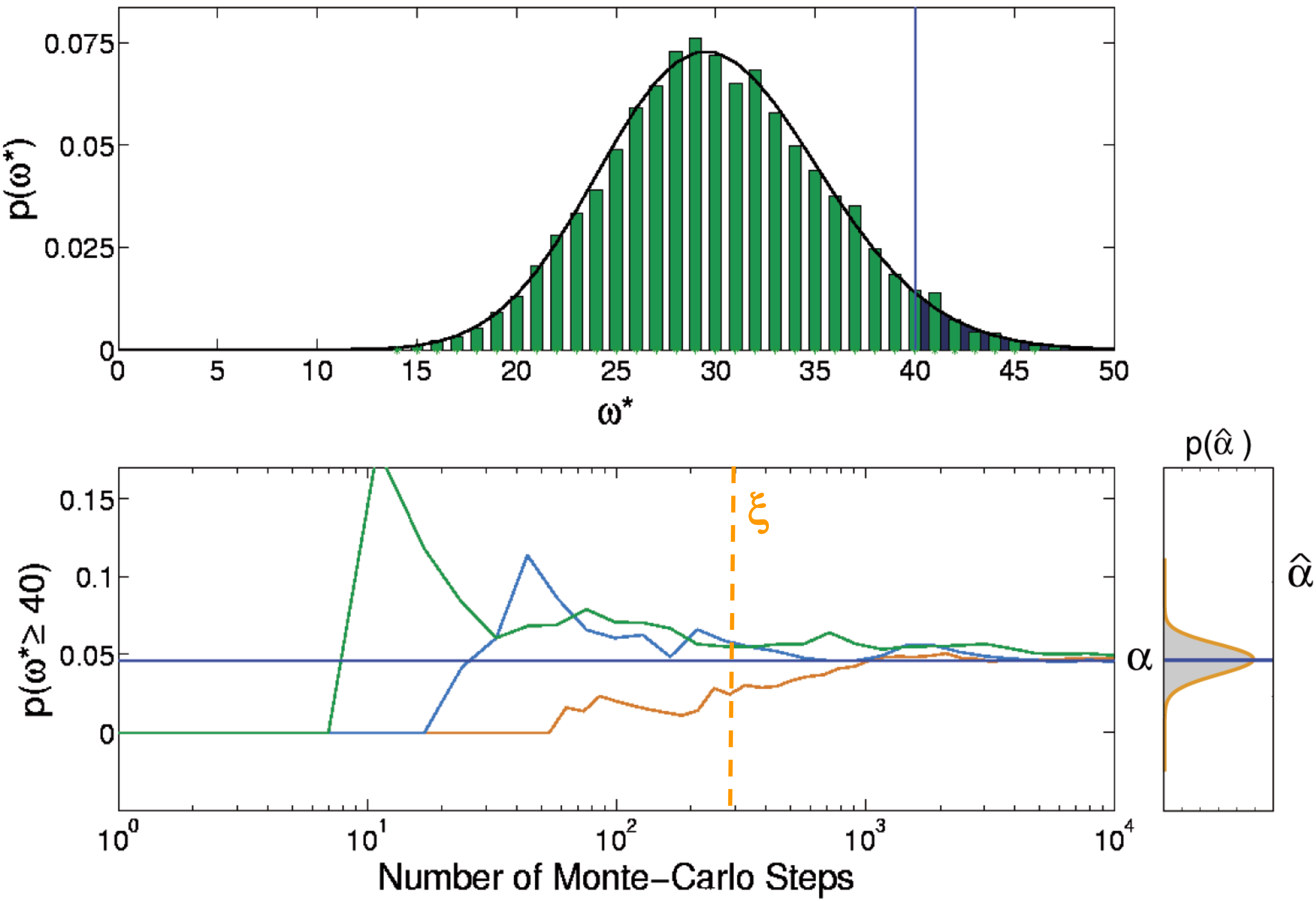
$$\hat{\alpha} = p\,(\omega^* \geq c) = \frac{\#(\omega^* \geq c)}{\xi}$$

Elfron and Tibshirani : 'An Introduction to the Bootstrap', Chapmann and Hall / CRC, 1998



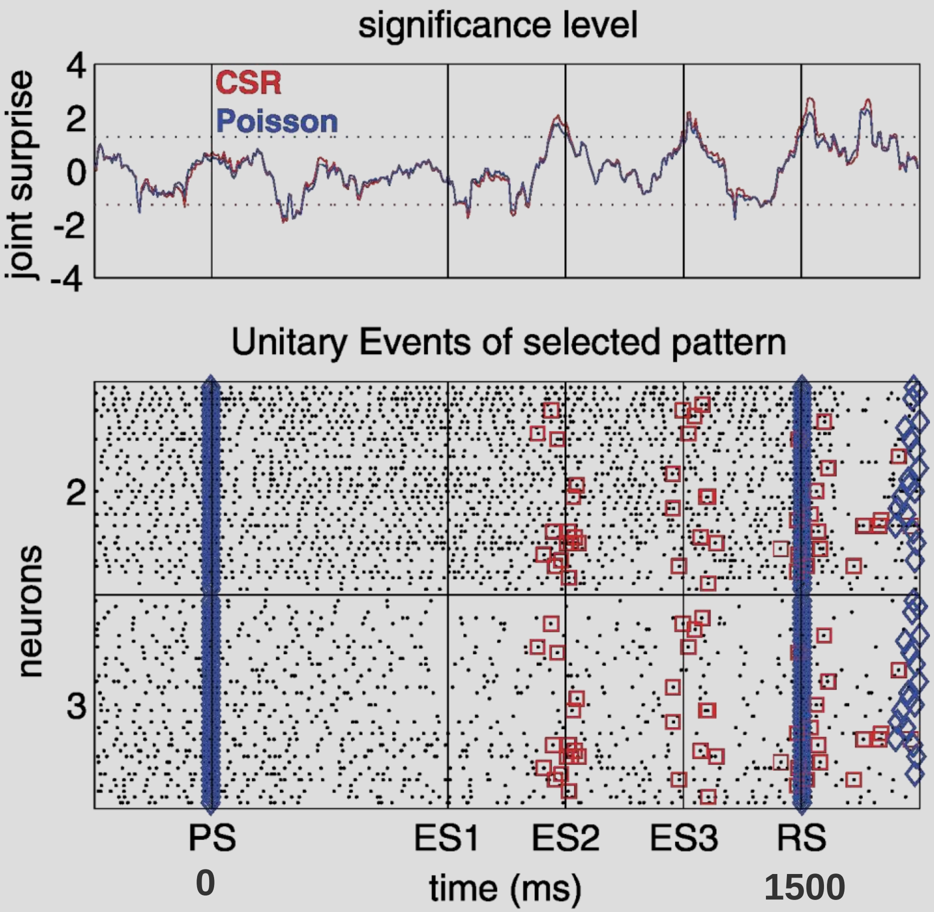
1
0.8
0.6
0.4
0.2
0
p(ω*)
0
5
10
15
20
25
30
35
40
45
50
ω*
0.15
0.1
0.05
0
p(ω* ≥ 40)
10^0
10^1
10^2
10^3
10^4
Number of Monte-Carlo Steps

Video : Resampling data class Ω_0



CSR-Method vs. Poisson based

Experimental data : CSR vs. Poisson based significance test



Data : monkey primary motor cortex
38 trials: delayed-pointing task
Riehle, Grün, Diesmann, Aertsen (1997) Science 278: 1950-1953

Analysis : coincidence window: 5 ms
test level: 5%
ξ=20000 ↔ afp= 0.05%

CSR Method : significant epochs are slightly more prominent