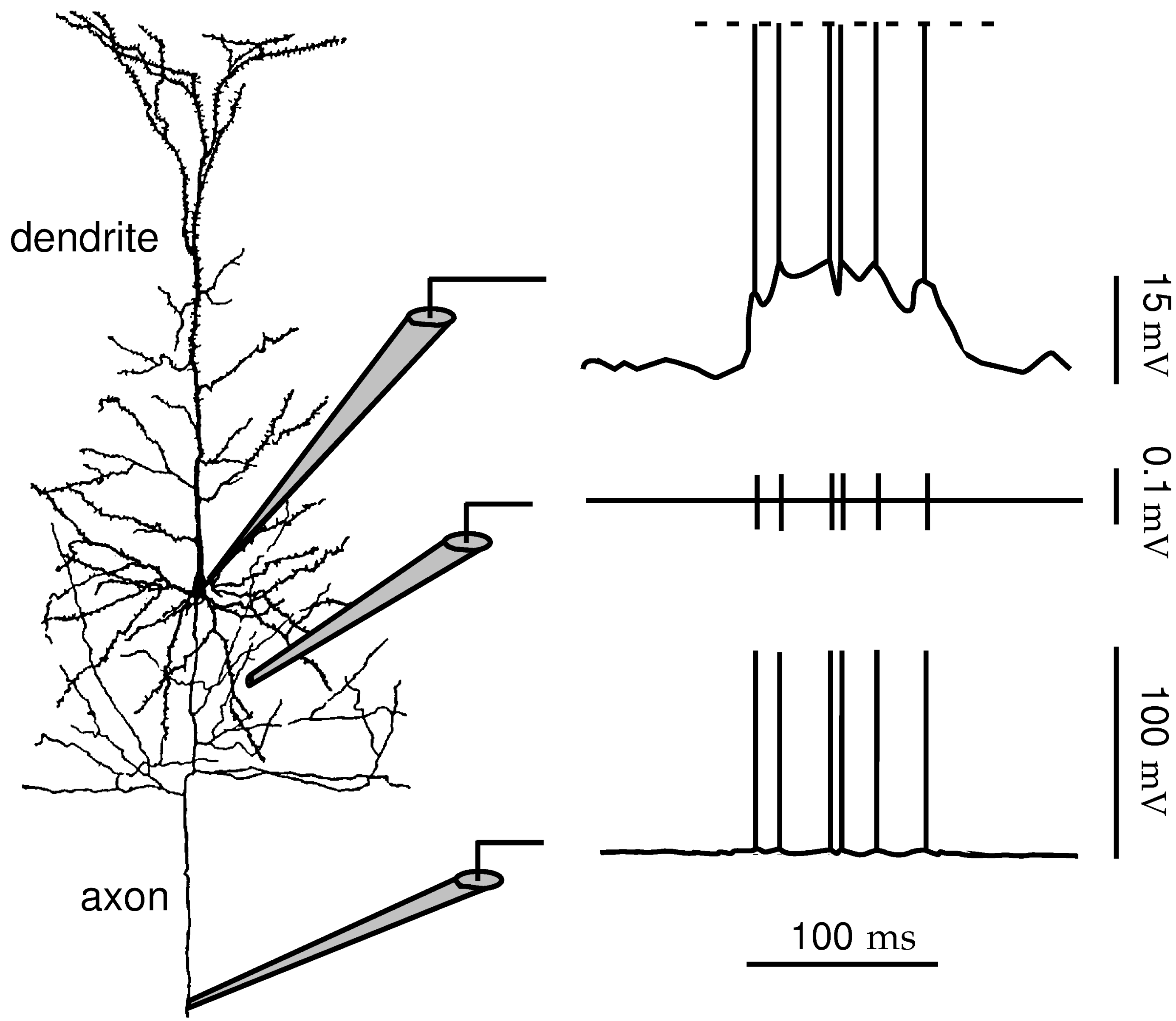
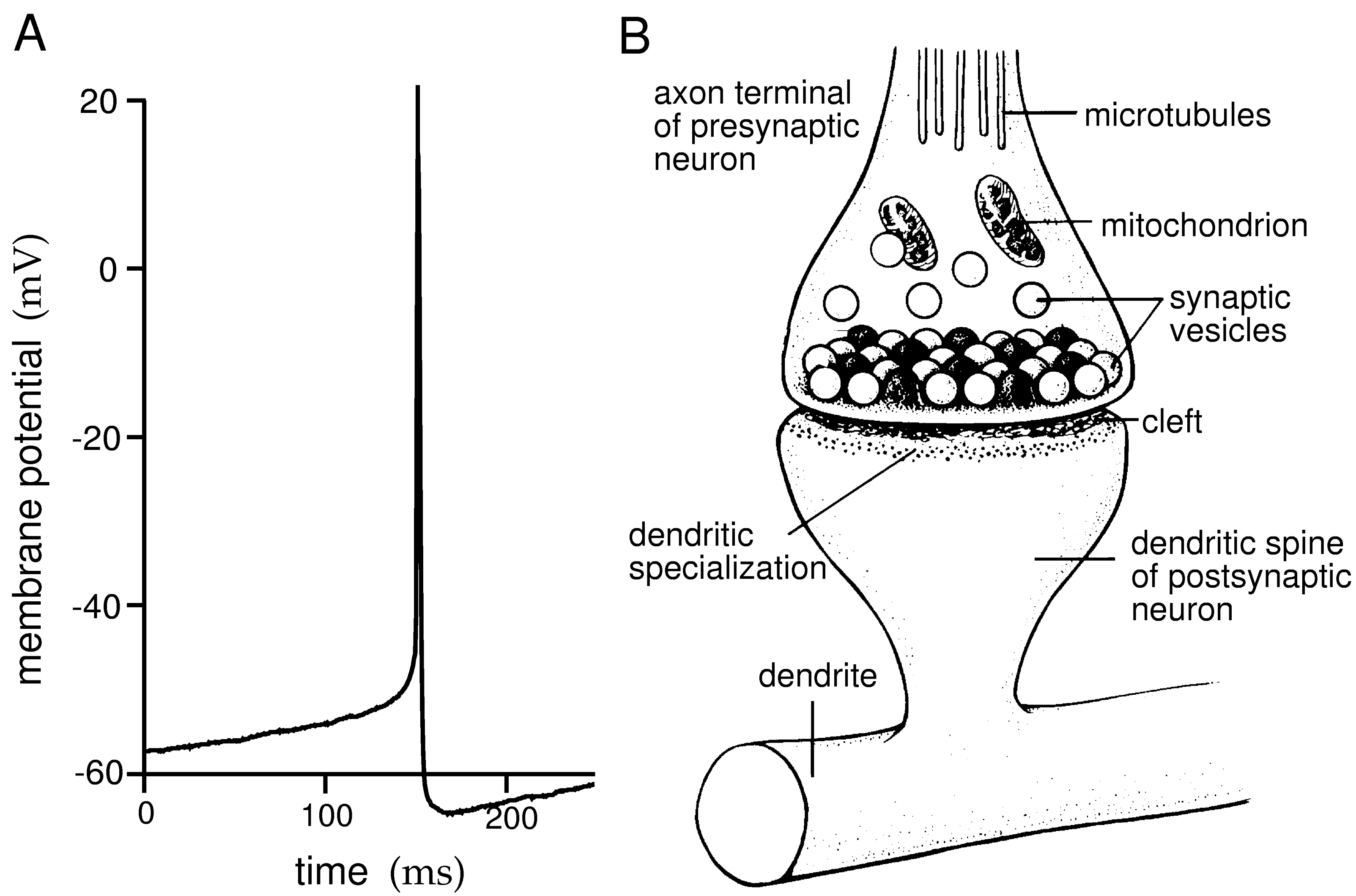
Lecture 3: Spike train statistics, part 1

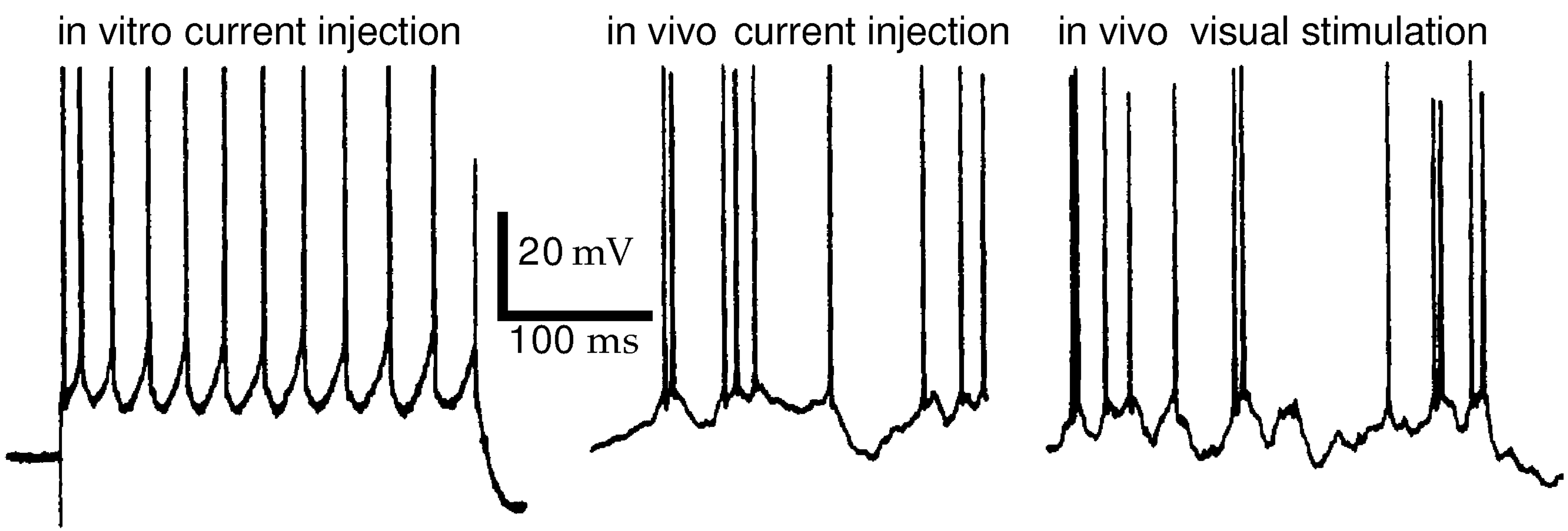


$$\mathcal{S} = \{t_1, t_2, \ldots t_N\}$$

Nothing but time matters:



Sources of 'noise' in neural recordings:



Point processes are a good start:

$$\mathcal{S} = \{t_1, t_2, \ldots t_N\}$$

$$s(t) = \sum_{i=1}^{N} \delta(t - t_i)$$

$$\begin{aligned} N(t) \text{ such that } & N(t) \geq 0 \\ & N(t + \Delta t) \geq N(t) \end{aligned}$$

The homogeneous Poisson process:

Recall that the Poisson distribution is a distribution on an integer random variable $n \geq 0$. If $n \sim \mathsf{Poiss}[\mu]$ then $\mathsf{P}[n] = \frac{\mu^n e^{-\mu}}{n!}$. The parameter μ is the mean of the distribution.

$$\mathcal{S} = \{t_1, t_2, \ldots t_N\}$$

- events occur with constant rate, rT (or μ)
- every sequence of n spikes has equal probability

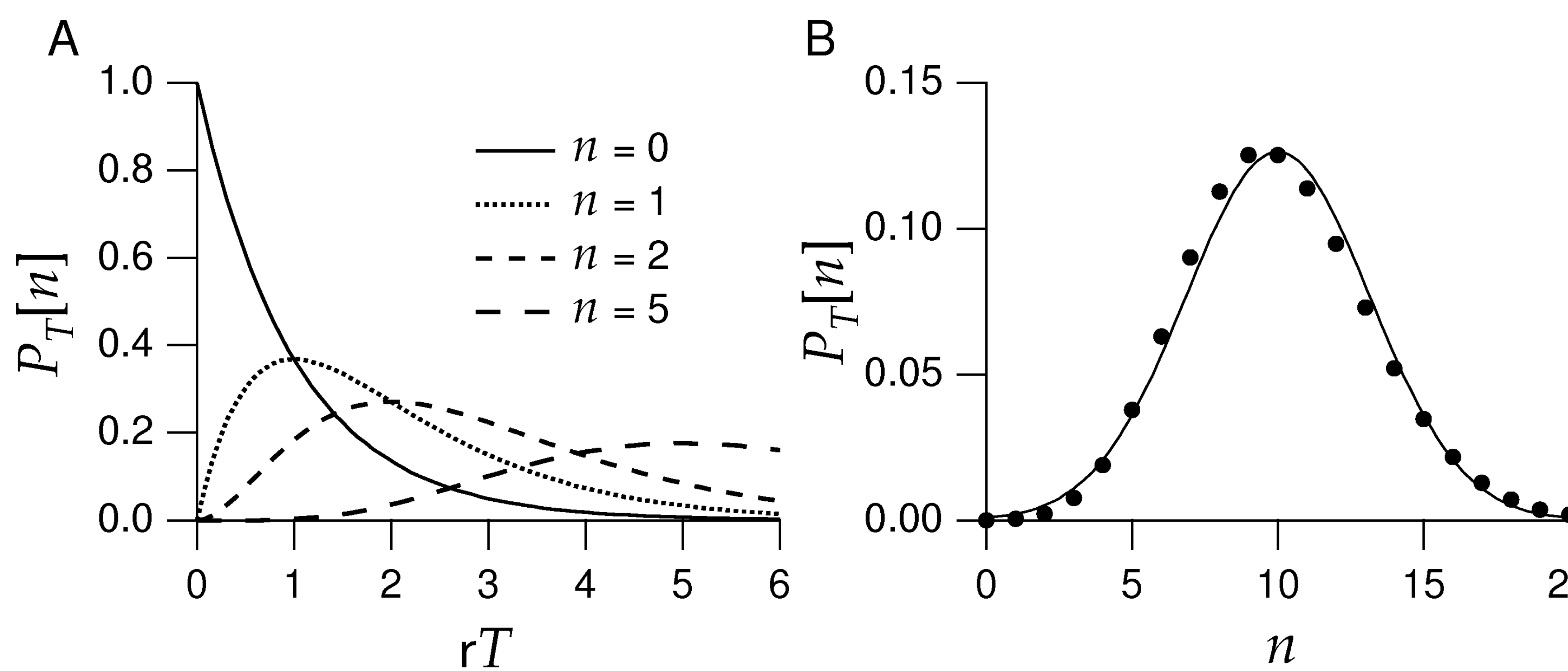
$$P[t_1, t_2, \ldots, t_n] = n! P_T[n] \left(\frac{\Delta t}{T}\right)^n$$

Probability of seeing n spikes in an interval:

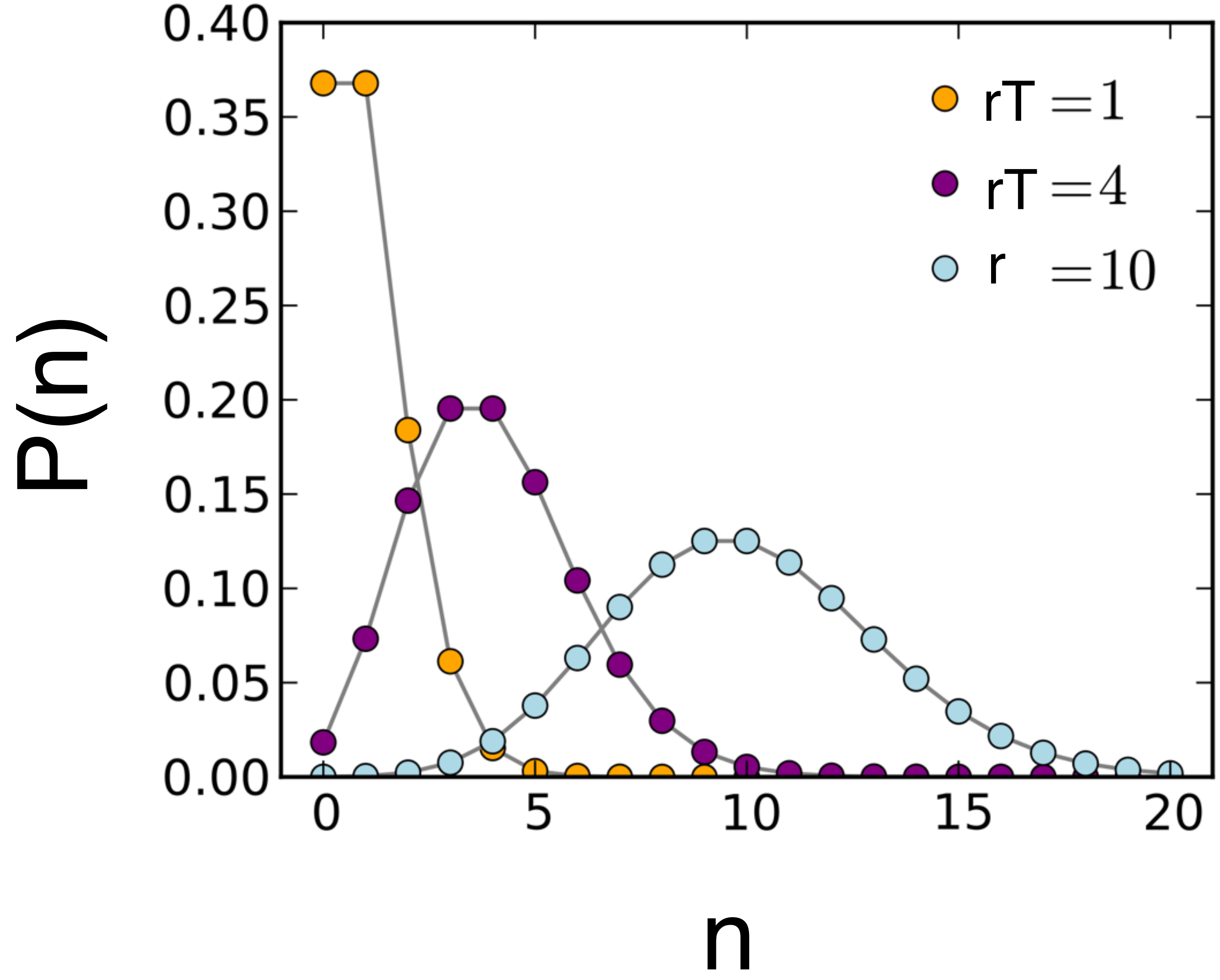
$$P_T[n] = \lim_{\Delta t \to 0} \frac{M!}{(M-n)!n!} (\mathrm{r}\Delta t)^n (1 - \mathrm{r}\Delta t)^{M-n}$$

$$\lim_{\Delta t \to 0} (1 - \mathrm{r}\Delta t)^{M-n} = \lim_{\epsilon \to 0} \left((1+\epsilon)^{1/\epsilon}\right)^{-\mathrm{r}T} = e^{-\mathrm{r}T} = \exp(-\mathrm{r}T)$$

$$P_T[n] = \frac{(\mathrm{r}T)^n}{n!} \exp(-\mathrm{r}T)$$



$$P_T[n] = \frac{(\mathrm{r}T)^n}{n!} \exp(-\mathrm{r}T)$$



deriving the P(n) for a Poisson process:

chalkboard interlude

Mean and variance of the count in a Poisson

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_T[n] = \sum_{n=0}^{\infty} \frac{n(\mathsf{r}T)^n}{n!} \exp(-\mathsf{r}T)$$

$$\sigma_n^2(T) = \sum_{n=1}^{\infty} n^2 P_T[n] - \langle n \rangle^2 = \sum_{n=1}^{\infty} \frac{n^2(\mathsf{r}T)^n}{n!} \exp(-\mathsf{r}T) - \langle n \rangle^2$$

chalkboard interlude

Fano Factor:

$$\sigma_n^2 / \langle n \rangle = 1$$

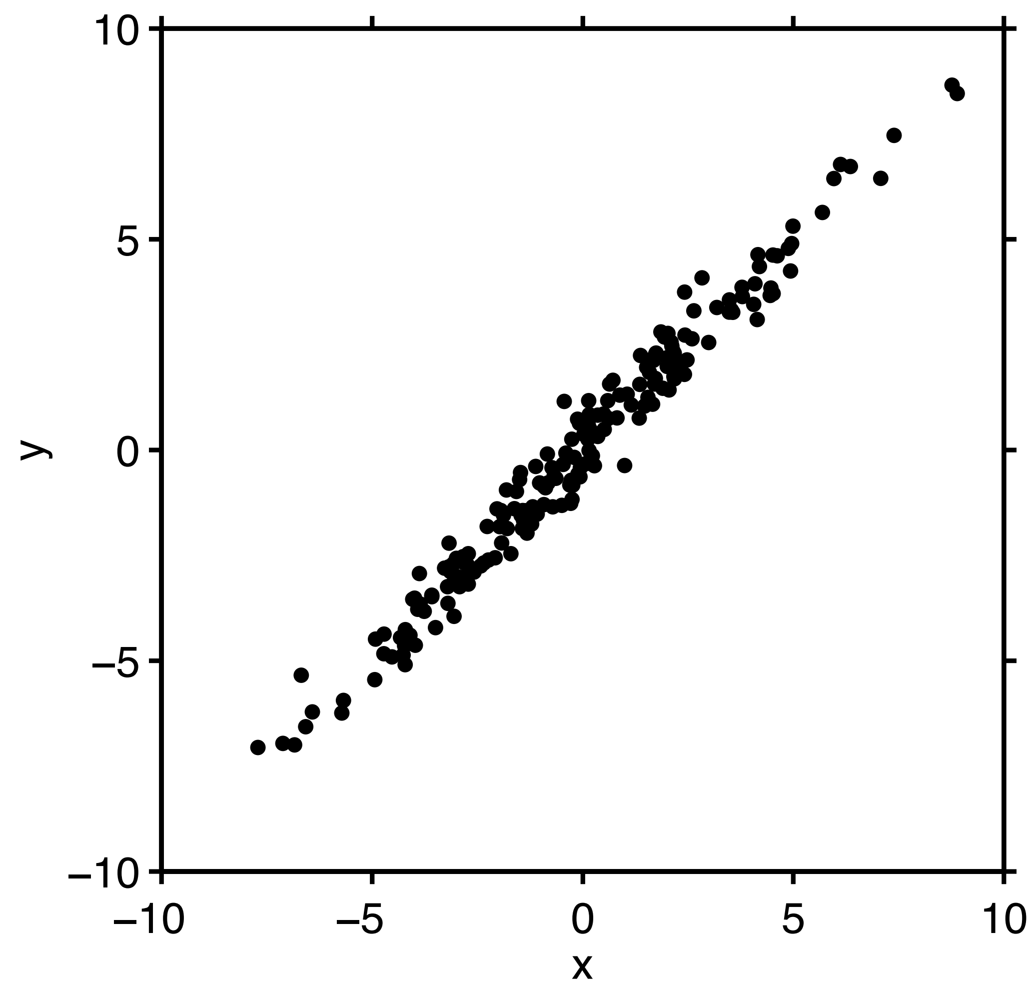
A brief introduction to information theory:

When is information theory useful? When you...

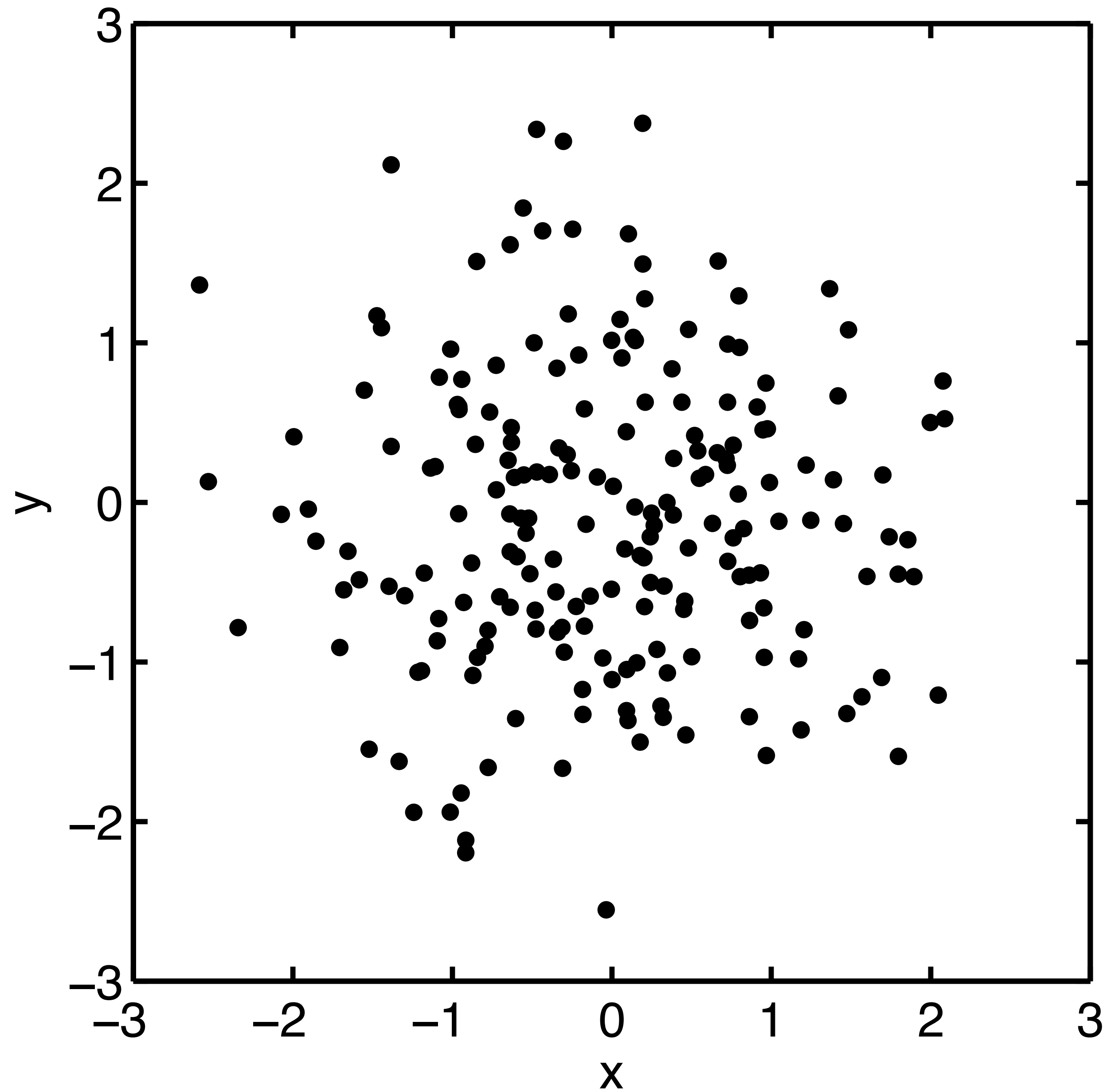
- **want to go beyond linear correlation**

- **have enough data to sample P(x,y)**

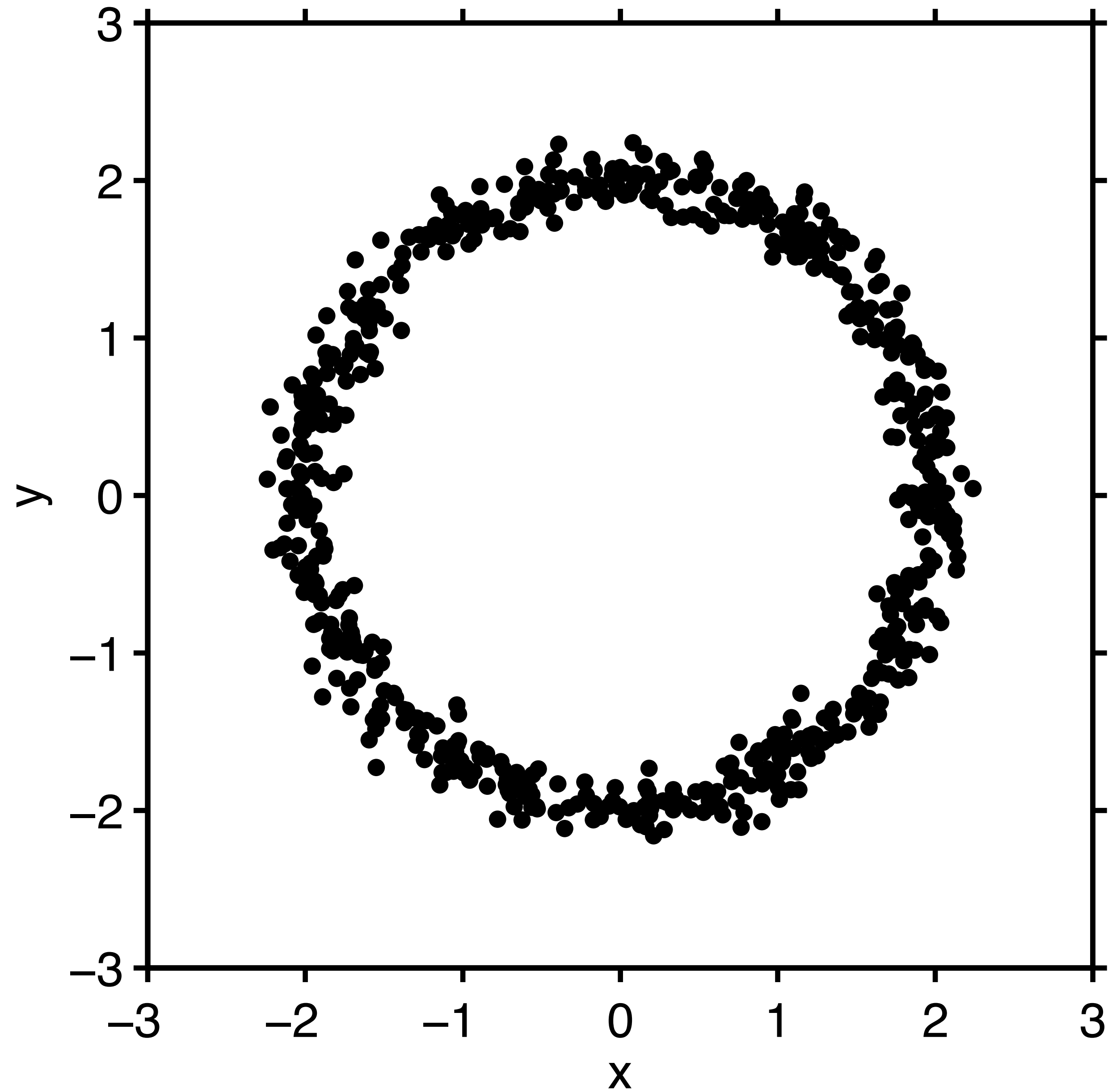
- **are not sure what your 'code' is**



Zero correlation, no information:



Zero correlation, but obvious information:



Entropy as a measure of uncertainty:

$$\begin{aligned} \text{uncertainty} &= \log(n) \\ &= \log(1/p) \\ &= -\log(p) \end{aligned}$$

$$u_i = -\log(p_i)$$

$$\langle u_i \rangle = -\sum_i p_i \log(p_i)$$

$$S(X) = -\sum_x p(x) \log_2(p(x))$$

Information == reduction in uncertainty

Mutual information:

$$
\begin{aligned}
I(A;B) &= S(A) - S(A|B) \\
&= S(B) - S(B|A) \\
&= \sum_{a,b} P(a,b) \log_2 \left(\frac{P(a,b)}{P(a)P(b)} \right) \\
&= \sum_{a,b} P(a)P(b|a) \log_2 \left(\frac{P(b|a)}{P(b)} \right)
\end{aligned}
$$

useful formulae

Product rule:

$$P(a,b) = P(a|b)P(b)$$

Sum rule:

$$\begin{aligned} P(a) &= \sum_b P(a,b) \\ &= \sum_b P(a|b)P(b) \end{aligned}$$

Bayes' rule:

$$\begin{aligned} P(a|b) &= \frac{P(b|a)P(a)}{P(b)} \\ &= \frac{P(b|a)P(a)}{\sum_{a'} P(b|a')P(a')} \end{aligned}$$

more useful formulae

Additivity:

$$S(A,B) = S(A) + S(B) \iff P(a,b) = P(a)P(b)$$

Chain rule:

$$S(A,B) = S(A) + S(B|A) = S(B) + S(A|B)$$

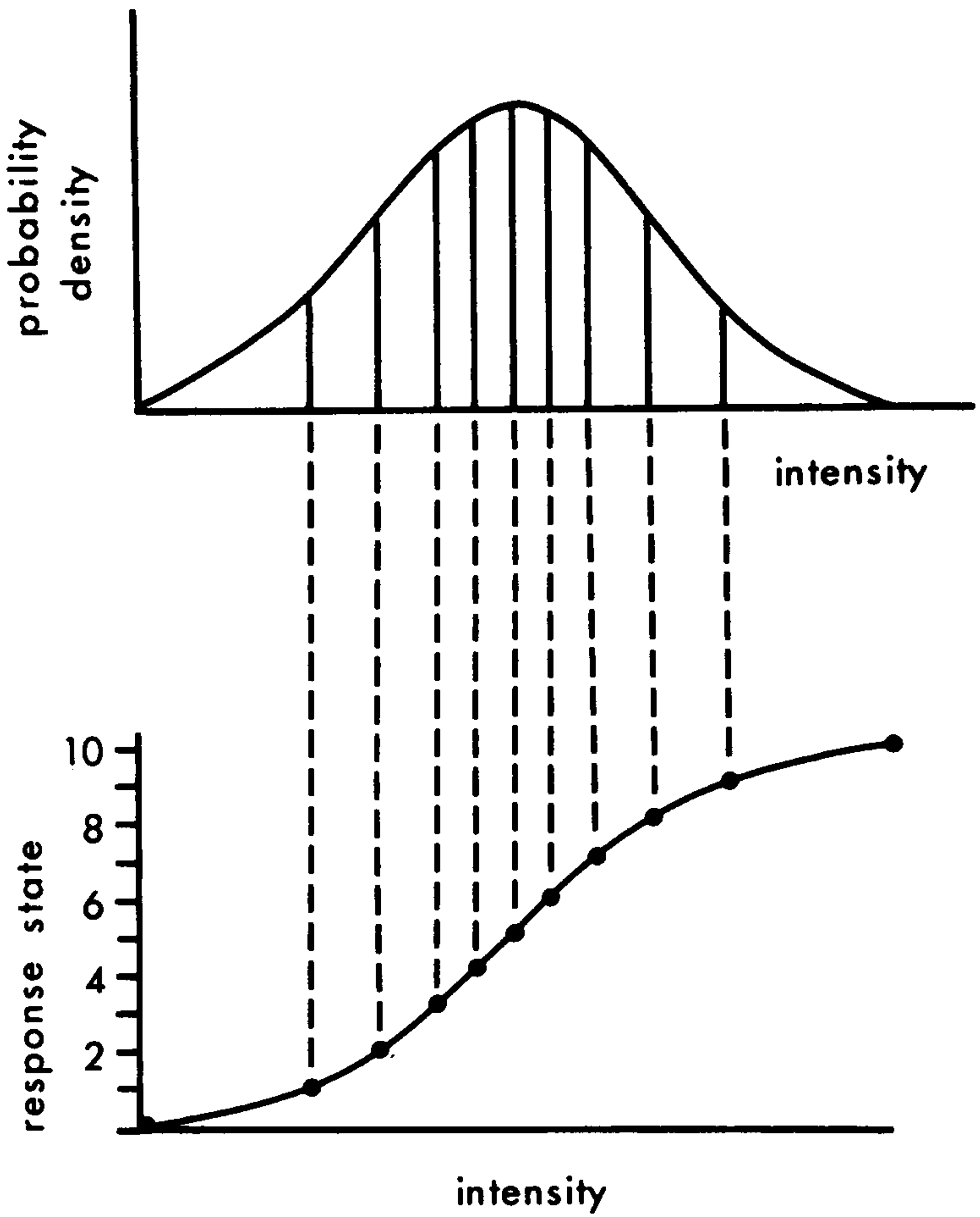
Kullback-Liebler divergence (D_{KL}):

$$D_{\mathrm{KL}}(P,Q) = \sum_a P(a)\log_2 \frac{P(a)}{Q(a)}$$

Mutual information ≥ 0:

$$I(A;B) = S(A) - S(A|B)$$

Efficient coding in single neurons



The efficient coding hypothesis, brief history:

- Claude Shannon (1948) *A Mathematical Theory of Communication*
- Fred Attneave (1954) *Some informational aspects of visual perception*
- Horace Barlow (1961) *Possible principles underlying the transformation of sensory messages*

Are sensory systems optimized for information transmission?

Information theory example, the weighing problem:

chalkboard interlude