PHY 341/641
Thermodynamics and Statistical Physics

Lecture 9

1. Probability concepts (Chapter 3 in STP)
   A. Bayes' theorm
   B. Binomial distribution

Bayes' theorm

Conditional probability --

Define $P(A|B) \equiv$ probability of A given B

Define $P(A|\overline{B}) \equiv$ probability of A given not B

$$\Rightarrow P(A) = P(A|B) + P(A|\overline{B})$$

$$\Rightarrow P(A \text{ and } B) = P(A|B)P(B) = P(B|A)P(A)$$

Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem continued:

Example: Suppose B represents a data set and A represents a model which can be generalized to several mutually exclusive models A_i

Generalized Bayes' theorem:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)}$$

If we assume that the set of A_i span all possible outcomes, then

$$P(B) = \sum_i P(B \mid A_i)P(A_i)$$

$$\Rightarrow P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_j P(B \mid A_j)P(A_j)}$$

Example 3.16. Bayes' theorem and the problem of false positives

Even though you have no symptoms, your doctor wishes to test you for a rare disease that only 1 in 10,000 people of your age contract. The test is 98% accurate, which means that, if you have the disease, 98% of the times the test will come out positive and 2% negative. Also, if you do not have the disease, the test will come out negative 98% of the time and positive 2% of the time. You take the test and it comes out positive. What is the probability that you have the disease? Answer the question before you read the solution using Bayes' theorem.

Let $P(D) \equiv$ probability of having the disease (0.0001 in this case)

$P(\overline{D}) \equiv P(N) \equiv$ probability of not having disease (0.9999)

$P(+|D) = 0.98 \qquad P(-|D) = 0.02$

$P(+|N) = 0.02 \qquad P(-|N) = 0.98$

Want to know $P(D|+) \equiv$ probability of having disease given positive test

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|N)P(N)}$$

$$= \frac{0.98 \cdot 0.0001}{0.98 \cdot 0.0001 + 0.02 \cdot 0.9999} = 0.0049 = 0.49\%$$

Compare with $P(D) = 0.0001 = 0.01\%$

Binomial distribution

Appropriate for describing systems with exactly two elemental outcomes such as

1. Coin Toss – HT -- P_H=p, P_T=1-p=q
2. Spin ↑↓ -- $P_\uparrow$=p, $P_\downarrow$=1- p=q

Typically, we are interested in the number n of elemental outcomes of type 1 (H, ↑, etc.) in a total of N processes.

Example: N=4, n=#H's, p=1/2=q

Configurations	n	P(n)
TTTT	0	1/16
HTTT,THTT,TTHT,TTTH	1	4/16
TTHH,THHT,HHTT,THTH,HTHT,HTTH	2	6/16
THHH,HTHH,HHTH,HHHT	3	4/16
HHHH	4	1/16

Binomial distribution

Example: N=4, n=#H’s, p=1/2=q

Configurations	n	P(n)
TTTT	0	1/16
HTTT,THTT,TTHT,TTTH	1	4/16
TTHH,THHT,HHTT,THTH,HTHT,HTTH	2	6/16
THHH,HTHH,HHTH,HHHT	3	4/16
HHHH	4	1/16

Example: N=3, n#H’s, p=1/2=q

Configurations	n	P(n)
TTT	0	1/8
HTT,THT,TTH	1	3/8
THH,HTH,HHT	2	3/8
HHH	3	1/8

Binomial distribution (continued)

In each of these cases,

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

Note that

$$\sum_{n=0}^{N} P_N(n) = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n q^{N-n} = 1$$

Because

$$1^N = (p+q)^N = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

Average value of n in binomial distribution

$$\langle n \rangle = \sum_{n=0}^{N} n P_N(n) = \sum_{n=0}^{N} n \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

$$\text{Note that}: \qquad n p^n q^{N-n} = p\left(\frac{\partial}{\partial p}\right)_q p^n q^{N-n}$$

$$\langle n \rangle = p\left(\frac{\partial}{\partial p}\right)_q \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n q^{N-n} = p\left(\frac{\partial}{\partial p}\right)_q (p+q)^N = pN$$

Similarly, we can show that:

$$\left\langle n^2 \right\rangle = \sum_{n=0}^{N} n^2 P_N(n) = \sum_{n=0}^{N} n^2 \frac{N!}{n!\left(N-n\right)!} p^n q^{N-n}$$

Note that: $$n^2 p^n q^{N-n} = \left(p\frac{\partial}{\partial p} \right)_q^2 p^n q^{N-n}$$

$$\left\langle n^2 \right\rangle = \left(p\frac{\partial}{\partial p} \right)_q^2 (p+q)^N = pN + p^2 N\left(N-1\right)$$

Variance:

$$\left\langle n^2 \right\rangle - \left\langle n \right\rangle^2 = pN + p^2 N\left(N-1\right) - p^2 N^2 = Np(1-p) = Npq$$

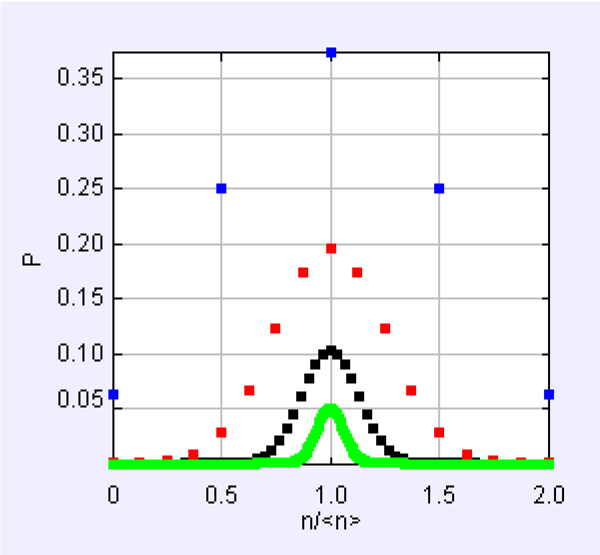
$$\left\langle n^2 \right\rangle - \left\langle n \right\rangle^2 \equiv \sigma^2 = Npq$$

Binomial distribution in the limit N →∞

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

$$P_{N\to\infty}(n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(n-\langle n\rangle)^2/2\sigma^2}$$

Result from stp_Binomial.jar



- N=4
- N=16
- N=60
- N=256

Stirling's approximation to factorial

For large m, Stirling's approximation to $m!$ is :

$$\ln m! \approx m \ln m - m + \frac{1}{2}\ln(2\pi m)$$

Evaluate Binomial distribution as Taylor's expansion of $\ln(P_N(n))$ about $\langle n \rangle$:

$$\ln P_N(n) \approx \ln P_N(\langle n \rangle) + (n - \langle n \rangle)\left.\frac{d \ln P_N(n)}{dn}\right|_{\langle n \rangle} + \frac{1}{2}(n - \langle n \rangle)^2 \left.\frac{d^2 \ln P_N(n)}{dn^2}\right|_{\langle n \rangle} \cdots$$

$$\ln A \qquad\qquad 0 \qquad\qquad -B$$

$$\Rightarrow P_N(n) \approx A e^{-\frac{B}{2}(n - \langle n \rangle)^2}$$

$$A \approx \frac{1}{\sqrt{2\pi\sigma^2}} \qquad B \approx \frac{1}{\sigma^2}$$

Summary:

$$P_N(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}$$

$$P_\infty(n) \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(n-\langle n \rangle)^2}{2\sigma^2}}$$

Poisson probability distribution –

Approximation to binomial distribution for small p

$$P_N(n) = \frac{N!}{n!\left(N-n\right)!} p^n q^{N-n}$$

$$\text{Assume} \quad N >> n \qquad\qquad p << 1$$

$$\frac{N!}{n!\left(N-n\right)!} \approx \frac{N^n}{n!}$$

$$\ln q^{N-n} = \left(N-n\right)\ln(1-p) \approx -\left(N-n\right)p \approx -Np$$

$$P_N(n) \approx \frac{N^n}{n!} p^n e^{-Np} = \frac{\left\langle n \right\rangle^n}{n!} e^{-\langle n \rangle}$$