Heiko Dumlich October 30, 2007

2 Task Theoretical Physics VI - Statistics

2.1 (non interacting spins)

The probability for a single spin to be up (q) or down (p) in a system of N spins without any external field or interaction between the spins is $p = q = 0.5$.

(a)

The probability $p_N(m)$ to have m spins up and $N-m$ spins down is taken out of the binomial distribution:

$$p_N(m) = \left(\begin{array}{c} N \\ m \end{array} \right) q^m (1-q)^{N-m} = \left(\begin{array}{c} N \\ m \end{array} \right) \left(\frac{1}{2}\right)^m \left(\frac{1}{2}\right)^{N-m} = \frac{N!}{(N-m)!m!} \left(\frac{1}{2}\right)^N$$

The normalisation condition is given using the binomial formula $\sum_{m=0}^{N} \left(\begin{array}{c} N \\ m \end{array} \right) q^m p^{N-m} = (q+p)^N$, with $p = (1-q)$:

$$\sum_{m=0}^{N} p_N(m) = \left(\frac{1}{2} + \frac{1}{2}\right)^N = (1)^N = 1$$

(b)

The mean and variance are meant to be calculated:

$$\begin{aligned}
\langle m \rangle &= \sum_{m=0}^{N} m P_N(m) \\
&= \sum_{m=0}^{N} m \frac{N!}{(N-m)!m!} q^m p^{N-m} \\
&= Nq \sum_{m=1}^{N} \frac{(N-1)!}{(N-m)!\,(m-1)!} q^{m-1} p^{(N-1)-(m-1)} \\
&= Nq \sum_{k=0}^{N-1} \frac{(N-1)!}{(N-k-1)!k!} q^k p^{(N-1)-k} \\
&= Nq \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} q^k p^{n-k} \\
&= Nq \sum_{k=0}^{n} P_n(k) \\
&= Nq
\end{aligned}$$

for our case with $q = \frac{1}{2}$ we therefore get:

$$\langle m \rangle = \frac{N}{2}$$

The alternative way uses the momentum generating function $M_m(t) := E(\exp(tm))$. Where $E(m) = \langle m \rangle$ means the expectation value which is equivalent to the mean value. Therefore we get:

$$M_m(t) = \langle \exp(tm) \rangle = \sum_{m=0}^{N} \binom{N}{m} q^m p^{N-m} e^{mt} = \sum_{m=0}^{N} \binom{N}{m} \left(qe^t\right)^m p^{N-m} = \left(qe^t + p\right)^N$$

The mean value then is

$$\langle m \rangle = \left.\frac{dM_m(t)}{dt}\right|_{t=0} = Nqe^t \left(qe^t + 1 - q\right)^{N-1}\Big|_{t=0} = Nq(1)^{N-1} = Nq$$

which leads to the same result:

$$\langle m \rangle = \frac{N}{2}$$

We know calculate the second derivative of the momentum generating function, which we will need to calculate the variance:

$$\begin{aligned}\left\langle m^2 \right\rangle &= \left.\frac{d^2 M_m(t)}{dt^2}\right|_{t=0} \\ &= N(N-1)q^2\left(qe^t + 1 - q\right)^{N-2} + Nqe^t\left(qe^t + 1 - q\right)^{N-1}\Big|_{t=0} \\ &= N(N-1)q^2 + Nq\end{aligned}$$

So in our case this is:

$$\left\langle m^2 \right\rangle = \frac{N}{4}(N+1)$$

With the definition of the variance:

$$\sigma^2 = M''(0) - \left[M'(0)\right]^2$$

we can easily calculate the variance:

$$\begin{aligned}\sigma^2 &= N(N-1)q^2 + Nq - N^2q^2 \\ &= N^2q^2 - Nq^2 + Nq - N^2q^2 \\ &= Nq(1-q)\end{aligned}$$

with $q = \frac{1}{2}$ this is in our case:

$$\sigma^2 = \frac{N}{4}$$

We now need to calculate the mean and variance for the dimensionless magnetization, which is defined by $M = 2m - N$, which can be easily done using the results we already obtained:

$$\langle M \rangle = \langle 2m - N \rangle = 2\langle m \rangle - N = 2\frac{N}{2} - N = 0$$

and the variance is:

$$\begin{aligned}\left\langle \Delta M^2 \right\rangle &= \left\langle M^2 \right\rangle - \left\langle M \right\rangle^2 \\ &= \left\langle (2m-N)^2 \right\rangle - 0 \\ &= \left\langle \left(4m^2 - 4mN + N^2\right) \right\rangle \\ &= 4\left\langle m^2 \right\rangle - 4\left\langle m \right\rangle N + N^2 \\ &= 4\frac{N}{4}(N+1) - 4\frac{N}{2}N + N^2 \\ &= N\end{aligned}$$

2.2 (typos in a book)

We are meant to find how many typos where not discovered in a book, where editor A found 200 and editor B found 150, while the were 100 typos found by both editors. This means:

$$\begin{aligned}N_A &= 200 \\ N_B &= 150 \\ N_{A\cap B} &= 100\end{aligned}$$

Using the formula for the conditional probability:

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

while A and B are independent, meaning $P(B|A) = P(B)$ we get:

$$P(B) = \frac{P(A\cap B)}{P(A)} \Leftrightarrow 1 = \frac{P(B)\,P(A)}{P(A\cap B)}$$

Now we use the frequency interpretation, which is valid for $N \gg 1$, which is fullfilled in our case:

$$\begin{aligned}P(A) &\approx \frac{N_A}{N} \\ P(B) &\approx \frac{N_B}{N} \\ P(A\cap B) &\approx \frac{N_{A\cap B}}{N}\end{aligned}$$

We also know the normalisation condition (for all mistakes N):

$$P(N) = 1 = \frac{N}{N}$$

Using the probability of all found mistakes by A and B:

$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

we can find the probability of the counter event with:

$$P(N) = P(A \cup B) + P\left(\overline{A \cup B}\right) \Leftrightarrow P\left(\overline{A \cup B}\right) = 1 - P(A \cup B)$$

Now inserting the terms we got from the frequency interpretation we get:

$$P\left(\overline{A \cup B}\right) = 1 - \frac{N_A}{N} - \frac{N_B}{N} + \frac{N_{A\cap B}}{N} = 1 - \frac{250}{N} = \frac{N_{\overline{A\cup B}}}{N} \Rightarrow N_{\overline{A\cup B}} = N - 250$$

Where $N_{\overline{A\cup B}}$ is the amount of missed typos. To find the total number of mistakes N we use:

$$1 = \frac{P(B)\,P(A)}{P(A \cap B)}$$

where $1 = P(N) = \frac{N}{N}$. Inserting the frequncy interpretation terms leads to:

$$\frac{N}{N} = \frac{\frac{N_B}{N}\frac{N_A}{N}}{\frac{N_{A\cap B}}{N}}$$

which can be transformed into:

$$N = \frac{N_B N_A}{N_{A\cap B}} = \frac{150 \cdot 200}{100} = 300$$

We insert the total number of mistakes in the formula for the amount of missed typos and we find:

$$N_{\overline{A\cup B}} = N - 250 = 300 - 250 = 50$$

So the number of missed typos is 50.

2.3 (probability and entropy)

We use the definition for the entropy of a probability distribution P according to $S = -k \sum P \ln P$. The definition of joint probability is given by:

$$P(A \cap B) = P(B|A) \cdot P(A) \Leftrightarrow P^{(A,B)}(i,j) = P^{(B)}(j|i) \cdot P^{(A)}(i)$$

with $i \in A$ and $j \in B$. We are meant to calculate the entropy $S = S_{(A,B)}$ and the distribution $P = P^{(A,B)}(i,j)$,

(a)

assuming strongly correlated events, which means:

$$P^{(B)}(j|i) = \delta_{ij}$$

for $i, j = 1, 2, \ldots, n$. Inserting this into the definition of P we find:

$$P = P^{(A,B)}(i,j) = \delta_{ij} \cdot P^{(A)}(i) = \delta_{ij} \cdot P^{(A)}(j)$$

This can be inserted in the definition of the entropy:

$$
\begin{aligned}
S(P) &= -k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\ln\left(\delta_{ij}\cdot P^{(A)}(j)\right)\\
&= -k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\left(\ln(\delta_{ij})+\ln\left(P^{(A)}(j)\right)\right)\\
&= -k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\ln(\delta_{ij})-k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\ln\left(P^{(A)}(j)\right)\\
&= -k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\ln(\delta_{ij})-k\sum_{j=1}^{n}P^{(A)}(j)\ln\left(P^{(A)}(j)\right)\\
&= -k\sum_{i,j=1}^{n}\delta_{ij}\cdot P^{(A)}(j)\ln(\delta_{ij})+S\left(P^{(A)}\right)
\end{aligned}
$$

Having a look at the first term we see the delta function in the ln-function, which for $i=j$ gives $\delta_{ii}=1$ leading to $\ln(1)=0$. Therefore we have to have a look at the case, where $i\neq j$, where $\delta_{ij}=0$, while $\ln(0)=?$. An easy way to find out, what the result is, is using a small variation $\lim_{m\to 0}\left(-kP^{(A)}(j)\ m\ln(m)\right)=\,?$. Using the rule of L'Hospital ($\frac{f(x)}{g(x)}=\frac{f'(x)}{g'(x)}$) we receive:

$$
\lim_{m\to 0}\left(\frac{-kP^{(A)}(j)}{\frac{1}{m}}\ln(m)\right)=\lim_{m\to 0}\left(kP^{(A)}(j)\,\frac{\frac{1}{m}}{\frac{1}{m^2}}\right)=\lim_{m\to 0}\left(kP^{(A)}(j)\ m\right)=0
$$

Therefore we can neglect the first term and for strong correlation we get the entropy:

$$
S(P)=S\left(P^{(A)}\right)
$$

Interpreting this, we see, that all events in B are events, that also lie in A. The entropy is therefore only decided by A.

(b)

assuming statistically independent events, which means:

$$
P^{(B)}(j|i)=P^{(B)}(j)
$$

which means:

$$
P^{(A,B)}(i,j)=P^{(B)}(j)\cdot P^{(A)}(i)
$$

Inserting this into the entropy definition we get:

$$
\begin{aligned}
S(P) &= -k\sum_{i,j=1}^{n}P^{(B)}(j)\cdot P^{(A)}(i)\ln\left(P^{(B)}(j)\cdot P^{(A)}(i)\right)\\
&= -k\sum_{i,j=1}^{n}P^{(B)}(j)\cdot P^{(A)}(i)\left(\ln\left(P^{(B)}(j)\right)+\ln\left(P^{(A)}(i)\right)\right)
\end{aligned}
$$

$$\begin{aligned}
&= -k\sum_{j=1}^{n} P^{(B)}(j)\ln\left(P^{(B)}(j)\right)\cdot\sum_{i=1}^{n} P^{(A)}(i) \\
&- k\sum_{i,j=1}^{n} P^{(A)}(i)\ln\left(P^{(A)}(i)\right)\cdot\sum_{j=1}^{n} P^{(B)}(j) \\
&= S\left(P^{(B)}\right)\cdot 1 + S\left(P^{(A)}\right)\cdot 1
\end{aligned}$$

This means the entropy is:

$$S(P) = S\left(P^{(A)}\right) + S\left(P^{(B)}\right)$$

The entropy in this case depends on A and B, resulting of the uncorrelated events.

(c)

The interpretation of the results from **(a)** and **(b)** yields the equation

$$S(P) = S\left(P^{(A,B)}\right) \leq S\left(P^{(A)}\right) + S\left(P^{(B)}\right)$$

which means, that the strongly correlated events yield a smaller entropy in every case then uncorrelated events, which yield the highest entropy in the case of no correlation.

(d)

To show that the entropy S of the distribution of $P = P(i) = \text{const}$ is the maximum for a uniform distribution, we have to have a look at S:

$$S = -k\sum_{i=1}^{N} P(i)\ln(P(i))$$

First we start out with the example, using $P(i) = \frac{1}{N}$ leads to:

$$S = -k\sum_{i=1}^{N}\frac{1}{N}\ln\left(\frac{1}{N}\right) = -k\sum_{i=1}^{N}\frac{1}{N}\left(\underbrace{\ln 1}_{=0} - \ln N\right) = k\sum_{i=1}^{N}\frac{1}{N}\ln N$$

If we execute the sum, we get:

$$S = k\frac{N}{N}\ln N = k\ln N$$

Now we have to test the maximum, where we can use our knowledge of analysis II. We use the lagrange multiplicator methode, where our boundary condition is $\sum_i P(i) = 1$, which leads to the boundary condition function $g(P(i)) = \sum_i P(i) - 1 = 0$. To find the extremum we use:

$$0 = \vec{\text{grad}}_P S(P(i)) - \lambda\vec{\text{grad}}_P g(P(i))$$

which leads to:

$$\begin{aligned}
0 &= -k\left(\ln\left(P\left(1\right)\right)+1\right)-\lambda \\
0 &= -k\left(\ln\left(P\left(2\right)\right)+1\right)-\lambda \\
\vdots &\ \vdots\ \vdots \\
0 &= -k\left(\ln\left(P\left(N\right)\right)+1\right)-\lambda
\end{aligned}$$

which can only be satisfied for $P(1) = P(2) = \ldots = P(N)$. Therefore the uniform distribution is needed for an extremum. To proof the maximum we now need to have a look at the 2nd derivative:

$$\Delta_P S\left(P\left(i\right)\right) = -k\frac{1}{P\left(i\right)}$$

Using the boundary condition, we find:

$$\sum_i P\left(i\right) = 1 = P\left(1\right) + \ldots + P\left(i\right) + \ldots + P\left(N\right) = N \cdot P\left(i\right) \Leftrightarrow P\left(i\right) = \frac{1}{N}$$

with $N > 0, N \in \mathbb{N}$. So the 2nd derivative is always negative with:

$$\Delta_P S\left(P\left(i\right)\right) = -kN$$

indicating a maximum. We can now calculated the lagrange multiplicator λ from the above equations:

$$\begin{aligned}
\lambda &= -k\left(\ln\left(\frac{1}{N}\right)+1\right) \\
\lambda &= k\ln N - k
\end{aligned}$$

2.4 (continuous random variable)

(a)

We are considering a particle performing a harmonic motion with $x(t) = x_0 \sin(\omega t + \phi)$, with unknown phase ϕ. Our interest lies in the speed $\frac{dx}{dt}$ at x. While this is a harmonic oscillator we can use Hooks law and therefore get the DGL:

$$F = m\frac{d^2x}{dt^2} = -Dx$$

which results to be the equation of the harmonic oscillator without damping. We can solve this one inserting $x(t)$, which leads to:

$$\omega = \sqrt{\frac{D}{m}}$$

Now we will have a look at the hamiltonian $H = T + V = E$ with $T = \frac{1}{2}m\left(\frac{dx}{dt}\right)^2$ and $V = \frac{D}{2}x^2$ which is the whole energy in the system:

$$H = E = \frac{1}{2}m\left(\frac{dx}{dt}\right)^2 + \frac{D}{2}x^2 \tag{1}$$

The whole energy in the system is constant. Having a look at the position $x\left(t=-\frac{\phi}{\omega}\right)=x_0$ leads to the energy:

$$\begin{aligned}E|_{x(t)=x} &= \left[\frac{1}{2}m\left(\omega x_0\cos\left(\omega t+\phi\right)\right)^2+\frac{D}{2}x_0^2\sin^2\left(\omega t+\phi\right)\right]_{t=-\frac{\phi}{\omega}} \\ &= \frac{1}{2}m\omega^2x_0^2\end{aligned}$$

This leads to a velocity at x of:

$$\begin{aligned}E|_{x(t)=x} &= \frac{1}{2}m\left(\frac{dx}{dt}\right)^2+\frac{D}{2}x^2 \\ \frac{1}{2}m\omega^2x_0^2 &= \frac{1}{2}m\left(\frac{dx}{dt}\right)^2+\frac{D}{2}x^2 \\ \frac{dx}{dt} &= \omega\sqrt{x_0^2-x^2}\end{aligned}$$

(b)

While the probability density $p(x)$ depends on the time the oscillator spends at x and the time it spends at x is inversely proportional to the speed (magnitude of velocity) at x we can write:

$$p(x)=c\frac{1}{\omega\sqrt{x_0^2-x^2}}$$

To estimate the proportionality constant c we can use the normalisation condition (the oscillation maximum is x_0 and the minimum $-x_0$, which results of the sin only giving values between 1 and -1):

$$\begin{aligned}\int_{-x_0}^{x_0}p(x)\,dx &= 1 \\ \frac{1}{c} &= \frac{1}{\omega}\int_{-x_0}^{x_0}\frac{1}{\sqrt{x_0^2-x^2}}dx \\ \frac{1}{c} &= \frac{1}{\omega x_0}\int_{-x_0}^{x_0}\frac{1}{\sqrt{1-\left(\frac{x}{x_0}\right)^2}}dx\end{aligned}$$

Using the substitution $\sin y=\frac{x}{x_0}$ with $\frac{dx}{dy}=x_0\cos y dy$ we get:

$$\begin{aligned}\frac{1}{c} &= \frac{1}{\omega x_0}\int_{\arcsin(-1)}^{\arcsin(1)}\frac{x_0\cos y}{\sqrt{1-\sin y^2}}dy \\ \frac{1}{c} &= \frac{1}{\omega}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{\cos y}{\cos y}dy \\ c &= \frac{\omega}{\pi}\end{aligned}$$

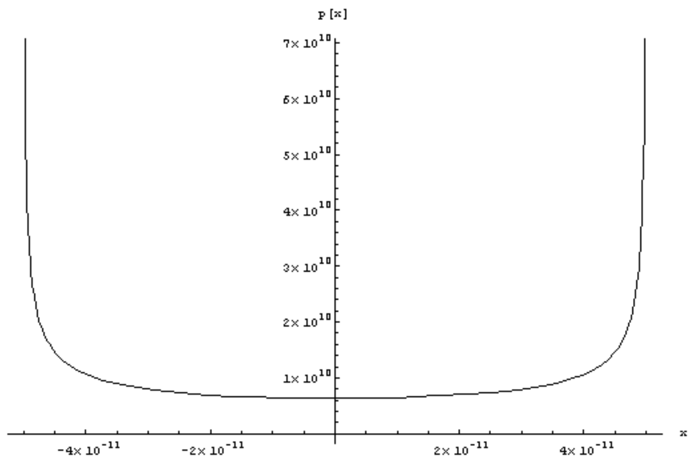

Figure 1: sketch of $p(x)$ with $x_0 = 0.5 \cdot 10^{-10}$, which is the bohr radius

Therfore the probability density is given with:

$$p(x) = \frac{1}{\pi\sqrt{x_0^2 - x^2}}$$

Which means, that the oscillator's probability density is independent of it's frequency ω (which by itself is dependent of the Hooks constant D and it's mass m).

(c)

From the sketch of $p(x)$ (fig. 1) we see the most probable value of $p(x)$ are $\pm x_0$ while $x = 0$ is the least probable. To get the mean value we calculate:

$$\begin{aligned} \langle x \rangle &= \int_{-x_0}^{x_0} x p(x)\, dx \\ &= \int_{-x_0}^{x_0} \frac{x}{\pi\sqrt{x_0^2 - x^2}} dx \\ &= \left[-\frac{1}{\pi}\sqrt{x_0^2 - x^2} \right]_{-x_0}^{x_0} \\ &= 0 \end{aligned}$$

which seems not making much sense, till $p(x) > 0$ at least at some points and it will never be negative, so we should get a positive mean value, at least the sketch tells us this. But if we think about $\langle x \rangle$ to be the position of the Oscillator, we know, that because of the symmetry 0 is the right answer.

Just some extra stuff I had a look at:

Now we want to have a look at what happens, when we have a look at $\langle p(x) \rangle$ (for the case that $f(p) = p(x)$, while $\langle p \rangle = \int_{-x_0}^{x_0} p\, f(p)\, dp$, which for sure isn't the case, but while we don't have $f(p)$ this is a first try):

$$\begin{aligned}
\langle p(x)\rangle &= \int_{-x_0}^{x_0} (p(x))^2 \; dx \\
&= \int_{-x_0}^{x_0} \frac{1}{\pi^2 (x_0^2 - x^2)} dx \\
&= \frac{1}{\pi^2 x_0^2} \int_{-x_0}^{x_0} \frac{1}{\left(1 - \left(\frac{x}{x_0}\right)^2\right)} dx
\end{aligned}$$

Using the substitution $\sin y = \frac{x}{x_0}$ with $\frac{dx}{dy} = x_0 \cos y dy$ we get:

$$\begin{aligned}
\langle p(x)\rangle &= \frac{1}{\pi^2 x_0^2} \int_{\arcsin(-1)}^{\arcsin(1)} \frac{x_0 \cos y}{(1 - \sin y^2)} dy \\
&= \frac{1}{\pi^2 x_0^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\cos y} dy \\
&= \frac{1}{\pi^2 x_0} \left[2\text{arctanh}\left(\tan\left(\frac{y}{2}\right)\right)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \frac{1}{\pi^2 x_0} \left[\infty - (-\infty)\right]
\end{aligned}$$

The integral doesn't converge unluckily.