PHYSICS 210A : EQUILIBRIUM STATISTICAL PHYSICS
HW ASSIGNMENT #1 SOLUTIONS

(1) Compute the information entropy in the Fall 2024Physics 140A grade distribution. See https://courses.physics.ucsd.edu/2024/Fall/physics140a .

Solution :

$\sum_n N_n = 59$	A+	A	A-	B+	B	B-	C+	C	C-	D	F
N_n	4	16	5	5	8	5	3	3	3	4	3
p_n	0.0678	0.271	0.0847	0.0847	0.136	0.0847	0.0508	0.0508	0.0508	0.0678	0.0508
$-p_n \log_2 p_n$	0.263	0.522	0.302	0.302	0.391	0.302	0.219	0.219	0.219	0.263	0.219

Table 1: F24 Physics 140A final grade distribution.

Assuming the only possible grades are A+, A, A-, B+, B, B-, C+, C, C-, D, F (11 possibilities), then from the chart we produce the entries in Tab. 1. We then find

$$S = -\sum_{n=1}^{11} p_n \log_2 p_n = 3.21 \text{ bits}$$

For maximum information, set $p_n = \frac{1}{11}$ for all n, whence $S_{\text{max}} = \log_2 11 = 3.46$ bits.

(2) Show that the Poisson distribution $P_\nu(n) = \frac{1}{n!}\,\nu^n\,e^{-\nu}$ for the discrete variable $n \in \mathbb{Z}_{\geq 0}$ tends to a Gaussian in the limit $\nu \to \infty$.

Solution:

For large fixed ν, $P_\nu(n)$ is maximized for $n \sim \nu$. We can see this from Stirling's asymptotic expression,

$$\ln n! = n \ln n - n + \tfrac{1}{2} \ln n + \tfrac{1}{2} \ln 2\pi + \mathcal{O}(1/n) \ ,$$

which yields

$$\ln P_\nu(n) = n \ln \nu - n \ln n - \nu + n - \tfrac{1}{2} \ln n - \tfrac{1}{2} \ln 2\pi$$

up to terms of order $1/n$, which we will drop. Varying with respect to n, which we can treat as continuous when it is very large, we find $n = \nu - \frac{1}{2} + \mathcal{O}(1/\nu)$. We therefore write $n = \nu + \frac{1}{2} + \varepsilon$ and expand in powers of ε. It is easier to expand in powers of $\tilde{\varepsilon} \equiv \varepsilon + \frac{1}{2}$, and since n is an integer anyway, this is really just as good. We have

$$\ln P_\nu(\nu + \tilde{\varepsilon}) = -(\nu + \tilde{\varepsilon}) \ln\left(1 + \frac{\tilde{\varepsilon}}{\nu}\right) + \tilde{\varepsilon} - \tfrac{1}{2} \ln(\nu + \tilde{\varepsilon}) - \tfrac{1}{2} \ln 2\pi \ .$$

Now expand, recalling $\ln(1+z) = z - \frac{1}{2} z^2 + \ldots$, and find

$$\ln P_\nu(\nu + \tilde{\varepsilon}) = -\frac{\tilde{\varepsilon}(1+\tilde{\varepsilon})}{2\nu} - \ln\sqrt{2\pi\nu} + \frac{\tilde{\varepsilon}^2}{4\nu^2} + \ldots$$

Since $\nu \to \infty$, the last term before the ellipses is negligible compared with the others, assuming $\tilde{\varepsilon} = \mathcal{O}(\nu^0)$. Thus,

$$P_\nu(n) \sim (2\pi\nu)^{-1/2} \exp\left\{ - \frac{\left(n - \nu + \frac{1}{2}\right)^2}{2\nu} \right\} ,$$

which is a Gaussian.

(3) Frequentist and Bayesian statistics can sometimes lead to different conclusions. You have a coin of unknown origin. You assume that flipping the coin is a Bernoulli process, *i.e.* the flips are independent and each flip has a probability p to end up heads and probability $1 - p$ to end up tails.

(a) You perform 14 flips of the coin and you observe the sequence {HHTHTHHHTTHHHH}. As a frequentist, what is your estimate of p?

(b) What is your frequentist estimate for the probability that the next two flips will each end up heads? If offered even odds, would you bet on this event?

(c) Now suppose you are a Bayesian. You view p as having its own distribution. The likelihood $f(\mathrm{data}|p)$ is still given by the Bernoulli distribution with the parameter p. For the prior $\pi(p)$, assume a Beta distribution,

$$\pi(p|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, p^{\alpha-1}\,(1-p)^{\beta-1} \quad .$$

where α and β are hyperparameters. Compute the posterior distribution $\pi(p|\mathrm{data}, \alpha, \beta)$.

(d) What is the posterior predictive probability $f(\mathsf{HH}\,|\,\mathrm{data}, \alpha, \beta)$?

(e) Since *a priori* we don't know anything about the coin, it seems sensible to choose $\alpha = \beta = 1$ initially, corresponding to a flat prior for p. What is the numerical value of the probability to get two heads in a row? Would you bet on it?

Solution:

(a) A frequentist would conclude $p = \frac{5}{7}$ since the trial produced ten heads and four tails.

(b) The frequentist reasons that the probability of two consecutive heads is $p^2 = \frac{25}{49}$. This is slightly larger than $\frac{1}{2}$, so the frequentist should bet! (Frequently, in fact.)

(c) Are you reading the lecture notes? You should, because this problem is solved there in §1.5.2. We have

$$\pi(p|\mathrm{data}, \alpha, \beta) = \frac{p^{9+\alpha}\,(1-p)^{3+\beta}}{\mathsf{B}(10+\alpha, 4+\beta)} ,$$

where the Beta function is $\mathsf{B}(\alpha, \beta) = \Gamma(\alpha)\,\Gamma(\beta)/\Gamma(\alpha+\beta)$.

(d) The posterior predictive is

$$p(\text{data}'|\text{data}) = \frac{\mathsf{B}(10+Y+\alpha, 4+M-Y+\beta)}{\mathsf{B}(10+\alpha, 4+\beta)} \,,$$

where Y is the total number of heads found among M tosses. We are asked to consider $M = 2$, $Y = 2$, so

$$f(\mathsf{HH}|\text{data}, \alpha, \beta) = \frac{\mathsf{B}(12+\alpha, 4+\beta)}{\mathsf{B}(10+\alpha, 4+\beta)} \,.$$

(e) With $\alpha = \beta = 1$, we have

$$f(\mathsf{HH}|\text{data}, \alpha, \beta)\Big|_{\alpha=\beta=1} = \frac{\mathsf{B}(13,5)}{\mathsf{B}(11,5)} = \frac{11 \cdot 12}{16 \cdot 17} = \frac{33}{68} = 0.4852941.$$

This is slightly less than $\frac{1}{2}$. Don't bet!

It is instructive to note that the Bayesian posterior prediction for a single head, assuming $\alpha = \beta = 1$, is

$$f(\mathsf{H}|\text{data}, \alpha, \beta)\Big|_{\alpha=\beta=1} = \frac{\mathsf{B}(11+\alpha, 4+\beta)}{\mathsf{B}(10+\alpha, 4+\beta)} = \frac{\mathsf{B}(12,5)}{\mathsf{B}(11,5)} = \frac{11}{16} \,.$$

The square of this number is $\frac{121}{256} = 0.4726565$, which is less than the posterior prediction for two consecutive heads, *even though our likelihood function is the Bernoulli distribution, which assumes the tosses are statistically independent*. The eager student should contemplate why this is the case.

(4) Professor Jones begins his academic career full of hope that his postdoctoral work, on relativistic corrections to the band structure of crystalline astatine under high pressure, will eventually be recognized with a Nobel Prize in Physics. Being of Bayesian convictions, Jones initially assumes he will win the prize with probability θ, where θ is uniformly distributed on $[0, 1]$ to reflect Jones' ignorance.

(a) After N years of failing to win the prize, compute Jones's chances to win in year $N + 1$ by performing a Bayesian update on his prior distribution.

(b) Jones' graduate student points out that Jones' prior is not parameterization-independent. He suggests Jones redo his calculations, assuming initially the Jeffreys prior for the Bernoulli process. What then are Jones' chances after his N year drought?

(c) Professor Smith, of the Economics Department, joined the faculty the same year as Jones. His graduate research, which concluded that poor people have less purchasing power than rich people, was recognized with a Nobel Prize in Economics[1] in his fifth year. Like Jones, Smith is a Bayesian, whose initial prior distribution was taken to be uniform. What is the probability he will win a second Nobel Prize in year 11? If instead Smith were a frequentist, how would he assess his chances in year 11?

[1]Strictly speaking, there is no such thing as a "Nobel Prize in Economics". Rather, there is a "Nobel Memorial Prize in Economic Sciences".

Solution:

(a) For the Beta distribution $\pi(\theta) = \theta^{\alpha-1}(1-\theta)^{\beta-1}/B(\alpha,\beta)$, one has

$$\langle\theta\rangle = \frac{\alpha}{\alpha+\beta} \quad .$$

Assuming $\alpha_0 = \beta_0 = 1$, under the Bayesian update rules, $\alpha_N = \alpha + P$ and $\beta_N = \beta + N - P$, where P is the number of successes in N years. Alas, for Jones $P = 0$, so $\alpha_N = 1$ and $\beta_N = N + 1$, meaning $f(\textsf{prize}|\textsf{reality}) = 1/(N+2)$.

(b) For the Jeffries prior, take $\alpha_0 = \beta_0 = \frac{1}{2}$, in which case $f(\textsf{prize}|\textsf{reality}) = 1/(2N+2)$.

(c) For Smith, we take $P = 1$ and $N = 10$, hence $f(\textsf{prize}|\textsf{reality}) = 2/(N+2) = \frac{1}{6}$. If Smith were a frequentist, he would estimate his chances at $p = \frac{1}{10}$.