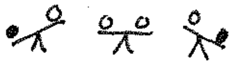
2.2 Two out of four heavier.

What is the uncertainty (entropy)?
What is a lower bound on the number of balance measurements needed? Is there such a procedure?

●●○○ Pick two heavier: can be done in $\binom{4}{2} = 6$ ways, all equally probable, so $P = \{p_i = \frac{1}{6}\}_{i=1}^{6}$.

$$\Rightarrow S = \sum_i p_i \log \frac{1}{p_i} = \log_2 6 \text{ bits or } \ln 6 \text{ nits}$$

Balance measurement ideally has equal probability for three outcomes  $\Rightarrow$ guarantees in that case $(\log_2 3)$ bits of info.

$2 \cdot \log_2 3 = \log_2 9$ bits $> S$, so at least two measurements are needed.

Cases Procedure: Conclusions:

A  $\xrightarrow{\text{switch the right one}}$  the two original were the light ones

B  $\xrightarrow{\text{switch right}}$  the two original were ~~light~~ heavy

C  $\xrightarrow{\text{switch right}}$  these are the two light ones

 the left one and the untested one are light

2.4 What is the relative info of two Gaussians with width (standard deviation) b_1, b_2?
Show that the quantity is nonnegative.

$$q(x) = \frac{1}{\sqrt{2\pi}\, b_1} \exp\left[-\frac{(x-\mu_1)^2}{2b_1^2}\right]$$

$$p(x) = \frac{1}{\sqrt{2\pi}\, b_2} \exp\left[-\frac{(x-\mu_2)^2}{2b_2^2}\right]$$

Going from q to p, the Kullback information is

$$K = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)}\, dx = \int_{-\infty}^{\infty} p(x) \Bigg(\underbrace{\ln \frac{b_1}{b_2}}_{①} \underbrace{- \frac{(x-\mu_2)^2}{2b_2^2}}_{②} + \underbrace{\frac{(x-\mu_1)^2}{2b_1^2}}_{③} \Bigg) dx$$

① $$\ln \frac{b_1}{b_2} \int_{-\infty}^{\infty} p(x)\, dx = \ln \frac{b_1}{b_2}$$

② $$-\frac{1}{2b_2^2} \int_{-\infty}^{\infty} p(x)\, (x-\mu_2)^2\, dx \underset{\substack{\text{definition}\\ \text{of variance}}}{=} -\frac{1}{2b_2^2} \cdot b_2^2 = -\frac{1}{2}$$

③ $$\frac{1}{2b_1^2} \int_{-\infty}^{\infty} p(x)\, (x-\mu_1)^2\, dx = \frac{1}{2b_1^2} \int_{-\infty}^{\infty} p(x) \left[(x-\mu_2)^2 + 2x\mu_2 - \mu_2^2 - 2x\mu_1 + \mu_1^2 \right] dx$$

$$= \frac{1}{2b_1^2} \left[\int_{-\infty}^{\infty} p(x)(x-\mu_2)^2 dx + (\mu_1^2 - \mu_2^2) \int_{-\infty}^{\infty} p(x)\, dx - 2(\mu_1 - \mu_2) \int_{-\infty}^{\infty} p(x)\, x\, dx \right]$$

$$= \frac{1}{2b_1^2} \left[b_2^2 + (\mu_1^2 - \mu_2^2) - 2\mu_2(\mu_1 - \mu_2) \right] = \frac{(\mu_1^2 - \mu_2^2) - 2(\mu_1 - \mu_2)\mu_2}{2b_1^2} + \frac{b_2^2}{2b_1^2}$$

$$= \frac{b_2^2}{2b_1^2} + \frac{(\mu_1 - \mu_2)(\mu_1 + \mu_2 - 2\mu_2)}{2b_1^2} = \frac{b_2^2}{2b_1^2} + \frac{(\Delta\mu)^2}{2b_1^2}$$

$$① + ② + ③ = \ln \frac{b_1}{b_2} - \frac{1}{2} + \frac{b_2^2}{2b_1^2} + \frac{(\Delta\mu)^2}{2b_1^2} \overset{?}{\geq} 0$$

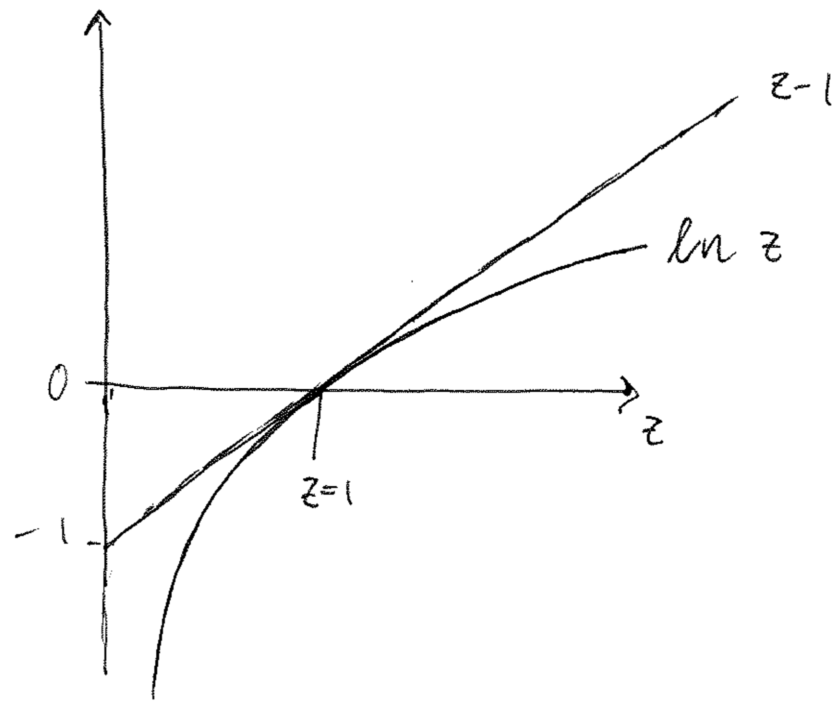
$\frac{(\Delta\mu)^2}{2b_1^2} \geq 0$, so we can just check if

$$\ln \frac{b_1}{b_2} - \frac{1}{2} + \frac{b_2^2}{2b_1^2} \geq 0.$$

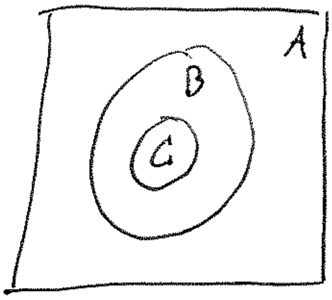
$$\Leftrightarrow \quad \frac{b_2^2}{b_1^2} - 1 \geq -2 \ln \frac{b_1}{b_2} = \ln \left(\frac{b_2}{b_1}\right)^2$$

Now let $\left(\frac{b_2}{b_1}\right)^2 = z$ and note that $z > 0$.

So, $z - 1 \overset{?}{\geq} \ln z$, $z > 0$. Yes! ■



2.6 Two-step observation.



We know that the result of an experiment is some element $k \in A$.

When we learn that $k \in B \subseteq A$,
which happens with probability $p_B = \mathbb{P}(k \in B)$,
the information gained is $I_{AB} = \log \frac{1}{p_B}$.

Then we learn that $k \in C \subseteq B$,
which happens with probability $\mathbb{P}(k \in C \mid k \in B) = \frac{p_C}{p_B}$,
so the additional information is

$$I_{BC} = \log \frac{1}{p_C/p_B}.$$

$$\begin{tikzcd}
 & k \in A \arrow[dl, "I_{AB}"'] \arrow[dd, "I_{AC}"] \\
k \in B \arrow[dr, "I_{BC}"'] & \\
 & k \in C
\end{tikzcd}$$

$$\Longrightarrow \quad I_{AB} + I_{BC} = \log \frac{1}{p_B} + \log \frac{p_B}{p_C} = \log \frac{1}{p_C}.$$

Had we instead learned immediately that $k \in C$,
the information would have been the same,

$$I_{AC} = \log \frac{1}{p_C}.$$

Example with 6-sided die:

$A = \{1,2,3,4,5,6\}$. Learn "k is odd" means $k \in B = \{1,3,5\}$.
Learn "k is 3" means $k \in C = \{3\}$.

$$I_{AB} = \log \frac{1}{3/6} = \log 2.$$
$$I_{AC} = \log \frac{1}{1/6} = \log 6.$$
$$I_{BC} = \log \frac{1}{1/3} = \log 3.$$

2.8 Radioactive decay.

p_τ = probability of not decaying before $t+\tau$ if it has not decayed at time t.

a) $$I = \ln \frac{1}{p_\tau}\Big|_{\tau=1/\lambda} = \ln \frac{1}{e^{-1}} = \ln e^{1} = 1 \text{ nit.}$$

b) $$I = \ln \frac{1}{p_\tau}\Big|_{\tau=1/\lambda} = \ln \frac{1}{1-e^{-1}} \text{ nits.}$$

2.9

$$\mathbb{E}[I_\tau] = p_\tau \ln \frac{1}{p_\tau} + (1-p_\tau) \ln \frac{1}{1-p_\tau}$$

Maximized when $p_\tau = \frac{1}{2}$, i.e. at the half-life time $T_{1/2}$: $p_{T_{1/2}} = \frac{1}{2}$.

$$p_\tau = \frac{1}{2} = e^{-\lambda\tau} \Rightarrow -\ln 2 = -\lambda\tau \Rightarrow \tau = \frac{\ln 2}{\lambda}$$

2.16 Maximize the entropy $S[P]$, $P = \{p_v\}_{v \in V}$, $V = \{-2, -1, 0, 1, 2\}$ with the requirements $\mathbb{E}[v] = 0$ and $\mathbb{E}[v^2] = 1$.

In other words, maximize $S[P] = \sum_{v \in V} p_v \ln \frac{1}{p_v}$

subject to
$$\sum_{v \in V} p_v v = 0,$$
$$\sum p_v v^2 = 1,$$
$$\sum p_v = 1.$$

Lagrangian:

$$L = S[P] + \lambda_1 \sum_v p_v v + \lambda_2 \left(\sum_v p_v v^2 - 1 \right) + \lambda_3 \left(\sum_v p_v - 1 \right)$$

$$0 = \frac{\partial L}{\partial p_v} = -1 - \ln p_v + \lambda_1 v + \lambda_2 v^2 + \lambda_3$$

$$\Rightarrow p_v = \exp\left[\lambda_3 - 1 + \lambda_1 v + \lambda_2 v^2 \right]$$

First, note that

$$\left.\begin{aligned} \frac{p_1}{p_{-1}} &= \exp\left[2\lambda_1\right] \\ \frac{p_2}{p_{-2}} &= \exp\left[4\lambda_1\right] \end{aligned}\right\}$$

so if we require $\lambda_1 \neq 0$, then $\mathbb{E}[v] = 2p_{-2}(e^{4\lambda_1} - 1) + p_1(e^{2\lambda_1} - 1) = 0$ only if $p_{-2} = p_{-1} = p_1 = p_2 = 0 \Rightarrow p_0 = 1$ $\Rightarrow \mathbb{E}[v^2] = 0$, which was not allowed. Therefore, $\lambda_1 = 0$.

Then, if $\lambda_1 = 0$, $p_1 = p_{-1}$ and $p_2 = p_{-2}$.

Now, require

$$\begin{cases} 0 = \frac{\partial L}{\partial \lambda_2} = \sum_v p_v v^2 - 1 \\ 0 = \frac{\partial L}{\partial \lambda_3} = \sum_v p_v - 1 \end{cases}$$

Rewrite using $\frac{p_2}{p_0} = \exp[4\lambda_2]$, $\frac{p_1}{p_0} = \exp[\lambda_2]$, $p_2 = p_{-2}$, $p_1 = p_{-1}$:

$$0 = \frac{\partial L}{\partial \lambda_2} = p_0\left(0 + 2e^{\lambda_2} + 2 \cdot 4\, e^{4\lambda_2}\right) - 1 \qquad (*)$$

$$0 = \frac{\partial L}{\partial \lambda_3} = p_0\left(1 + 2e^{\lambda_2} + 2e^{4\lambda_2}\right) - 1 \qquad (**)$$

$$\Rightarrow\ 0 = (*) - (**) = p_0\left(6e^{4\lambda_2} - 1\right)$$

We see in $(*)$ that $p_0 \neq 0$, so we must have $\left(6e^{4\lambda_2} - 1\right) = 0 \iff e^{\lambda_2} = 6^{-1/4}$.

$$(*) \Rightarrow p_0 = \frac{1}{2e^{\lambda_2} + 8e^{4\lambda_2}} = \frac{1}{2 \cdot 6^{-1/4} + 8 \cdot \frac{1}{6}} = \frac{3}{6^{3/4} + 4}$$

$$p_1 = p_0 e^{\lambda_2} = \frac{3}{6 + 4 \cdot 6^{1/4}}$$

$$p_2 = p_0 e^{4\lambda_2} = \frac{1}{2 \cdot 6^{3/4} + 8}$$