Mathematical Physics 202's Big Book of Problems

and solutions

Jonas Rademacker

Book of weekly problems for Mathematical Physics 202, University of Bristol.

Contents

Chapter 1

Practice Problems Week 1

1.1 Math basics

1.1.1 Derivatives

a) $f(x) = (\cos x - \sin x) \cdot e^x$. Show that $\frac{df}{dx} = -2\,\sin x\; e^x$
Use product rule: $(uv)' = (u')v + u(v')$ *with* $u = (\cos x - \sin x)$ *and* $v = e^x$*. The derivatives of* u *and* v*:*
$u' = (-\sin x - \cos x)$ *and* $v' = e^x$*. Plugging this in:*
$\frac{df}{dx} = (-\sin x - \cos x)\, e^x + (\cos x - \sin x)\, e^x = -2 \sin x\, e^x$.

b) $g(t) = 2e^{at} \arcsin t$. What is $\frac{dg}{dt}$?
$2e^{ax}\left(a \cdot \arcsin x + \frac{1}{\sqrt{1-x^2}}\right)$

c) $g(t) = 2e^{at} \arcsin t$. What is $\frac{dg}{dt}(0)$?
2

d) $g(t) = 2e^{at} \arcsin t$. What is $\frac{dg}{da}$?
$2te^{at} \arcsin t$

e) $y(\theta) = \frac{\ln \theta}{\sqrt{\theta}}$ What is $\frac{dy}{d\theta}$?
$\frac{2-\ln\theta}{2\theta\sqrt{\theta}}$

f) $y(x) = \ln \sqrt{4x - x^2}$. Calculate y'.
$y = \ln\left(4x - x^2\right)^{\frac{1}{2}} = \frac{1}{2} \ln\left(4x - x^2\right)$
$y' = \frac{2-x}{4x-x^2}$

1.1.2 Integrals

Remember the difference between indefinite integrals (=anti derivatives) and definite integrals:

- Calculate the indefinite integral $\int x\, dx$. Answer: $\frac{1}{2}x^2 + C$. Note that the notation $\int f(x)dx$ sometimes means the indefinite integral / anti-derivative of $f(x)$, and sometimes means the (definite) integral over all values, i.e. usually $-\infty$ and $+\infty$ (for angles $0, \ldots, 2\pi$).
- Calculate $\int_a^b x\, dx$. Answer: $\frac{1}{2}b^2 - \frac{1}{2}a^2$ (definite integral)
- Calculate $\int_0^2 x\, dx$. Answer: 2 (definite integral)

a) Calculate the indefinite integral $I = \int (-2\sin x\, e^x)\, dx$.
Follows directly from question (a) on derivatives: $I = (\cos x - \sin x) \cdot e^x + C$

b) Calculate $\int_0^1 \frac{x}{(1+x^2)^2}\, dx$ Hint: Substitute $u = 1 + x^2$.
$\frac{du}{dx} = 2x, \quad dx = \frac{du}{2x}$

$$\begin{aligned} I &= \int_0^1 \frac{x}{(1+x^2)^2} \\ &= \int_{u=1}^{2} \frac{x}{u^2} \frac{du}{2x} \\ &= \frac{1}{2} \int_1^2 \frac{1}{u^2} \\ &= \frac{1}{2} \left[-\frac{1}{u}\right]_1^2 = \frac{1}{4} \end{aligned}$$

c) Calculate $\int \frac{e^{2x}}{1+e^x} dx$. (Hint: substitute $u = 1 + e^x$)
$\frac{du}{dx} = e^x,\ dx = \frac{1}{e^x}\, dx$
With this: $I = \int \frac{e^x}{u}\, du = \int \frac{u-1}{u}\, du = \int \left(1 - \frac{1}{u}\right)\, du = u - \ln|u| + C$
Substitute back: $u = 1 + e^x$
$I = (1 + e^x) - \ln|1 + e^x| + C = e^x - \ln(1 + e^x) + K$ *with* $K = 1 + C$

d) Calculate $I = \int (1 + 2x)\, e^{-x}\, dx$. Hint: Use integration by parts: $\int u \cdot v'\, dx = u \cdot v - \int u' \cdot v\, dx$
Use $u = (1 + 2x),\ v' = e^{-x}$*, which gives* $v = -e^{-x}$ *and* $u' = 2$*. Hence*
$I = -\int (1 + 2x)\, e^{-x}\, dx = (1 + 2x)(-e^{-x}) - \int 2(-e^{-x})dx = -(3 + 2x) \cdot e^{-x} + C.$

e) Calculate $I = \int_{x=0}^{1} \int_{y=0}^{-x+1} x^2 y^2 dy\, dx$
Step one: integrate inner integral (y)*:*

$$\begin{aligned} \int_{y=0}^{-x+1} x^2 y^2 dy &= x^2 \int_{y=0}^{-x+1} y^2 dy \\ &= x^2 \left[\frac{1}{3} y^3\right]_{y=0}^{-x+1} \\ &= x^2 \left(\frac{1}{3}(-x+1)^3 - 0\right) \\ &= \frac{1}{3} x^2 (-x+1)^3 \\ &= \frac{1}{3}\left(-x^5 + 3x^4 - 3x^3 + x^2\right) \end{aligned}$$

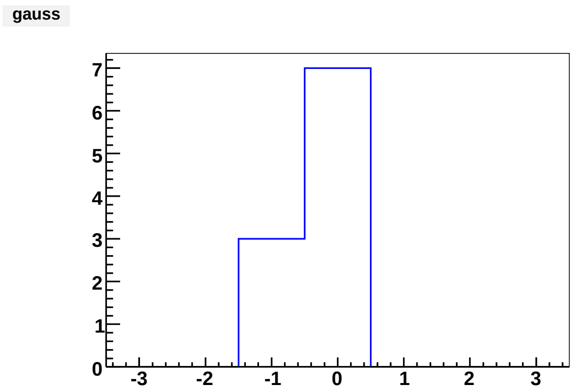
Now the outer integral (x):

$$
\begin{aligned}
I &= \frac{1}{3}\int_0^1 (-x^5 + 3x^4 - 3x^3 + x^2) \\
&= \frac{1}{3}\int \left[\frac{1}{6}(-x^6) + \frac{1}{5}\cdot 3x^5 \frac{1}{4}\cdot(-3)x^4 + \frac{1}{3}\cdot x^3\right]_0^1 \\
&= \frac{1}{180}
\end{aligned}
$$

1.2 Statistics

1.2.1 Means, variances, width

1) Calculate the mean, variance and standard deviation of the following histogram:



- *mean:*

$$
\begin{aligned}
\overline{x} &= \frac{1}{N_{\text{events}}}\sum_{i=0}^{N_{\text{events}}} x_i \quad \textit{(for individual events with values } \{x_i\}\textit{)} \\
&\approx \frac{1}{N_{\text{events}}}\sum_{j=0}^{N_{\text{bins}}} f_j \cdot x_{\text{bin j}} \quad \textit{(for binned data, where } f_j \textit{ is the number of events in the bin with center } x_{\text{bin j}}\textit{)} \\
&= \frac{1}{10}\cdot(7\cdot 0 + 3\cdot(-1)) \\
&= -0.3
\end{aligned}
$$

- *variance:*

$$
\begin{aligned}
V &= \frac{1}{N_{\text{events}}}\sum_{i=0}^{N_{\text{events}}} (x_i - \overline{x})^2 \quad \textit{(for individual events with values } \{x_i\}\textit{)} \\
&\approx \frac{1}{N_{\text{events}}}\sum_{j=0}^{N_{\text{bins}}} f_j \cdot (x_{\text{bin j}} - \overline{x})^2 \quad \textit{(for binned data, where } f_j \textit{ is the number of events in the bin with center } x_{\text{bin j}}\textit{)} \\
&= \frac{1}{10}\cdot\left(7\cdot((-0.3)-0)^2 + 3\cdot((-0.3)-(-1))^2\right) \\
&= 0.21
\end{aligned}
$$

- *standard deviation:*

$$\sigma = \sqrt{V} = 0.46$$

2) What is the 3rd central moment of the histogram in the prev. question? What would it be for a completely symmetric distribution?

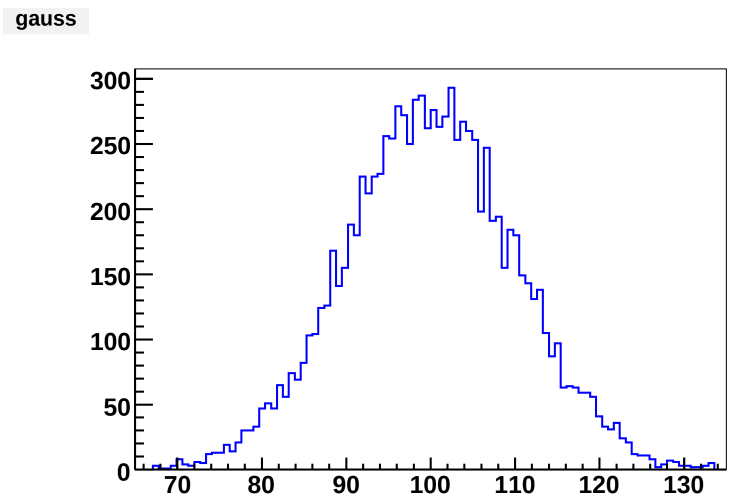
$$\begin{aligned} c_3 &= \frac{1}{N_{\text{events}}} \sum_{i=0}^{N_{\text{events}}} (x_i - \overline{x})^3 \quad \textit{(for individual events with values } \{x_i\}\textit{)} \\ &\approx \frac{1}{N_{\text{events}}} \sum_{j=0}^{N_{\text{bins}}} f_j \cdot (x_{\text{bin j}} - \overline{x})^3 \quad \textit{(for binned data, where } f_j \textit{ is the number of events in the bin with center } x_{\text{bin j}}\textit{)} \\ &= \frac{1}{10} \cdot \left(7 \cdot (0 - (-0.3))^3 + 3 \cdot ((-1) - (-0.3))^3\right) \\ &= \frac{1}{10} \cdot \left(7 \cdot (0.3)^3 + 3 \cdot (-0.7)^3\right) \\ &= -0.084 \end{aligned}$$

for a symmetric distribution, all odd moments (including this one) would be 0.

1.2.2 Standard Deviation and FWHF

Estimate the standard deviation σ of the following histogram, assuming it is roughly Gaussian (as many distributions in practice are), and using

$$FWHM \approx 2.35\sigma.$$



Generated with $\sigma = 10$.

1.2.3 Means, variances, standard deviations, covariances

For the following datasets, calculate the mean and standard deviation of each, x and y. Also calculate the covariance and the correlation coefficient.

a)

x	2	4	12
y	4	8	24

- *The means:*

$$\overline{x} = (2+4+12)/3 = 18/3 = 6$$
$$\overline{y} = (4+8+24)/3 = 12$$

- *The standard deviations* $\sigma = \sqrt{V}$*:*

$$\begin{aligned}
V_x &= ((2-6)^2 + (4-6)^2 + (12-6)^2)/3 = \frac{164}{3} 18\frac{2}{3} \approx 18.667 \\
\sigma_x &= \sqrt{V_x} = 4.32 \\
V_y &= ((4-12)^2 + (8-12)^2 + (14-12)^2)/3 = \frac{224}{3} = 74\frac{2}{3} \approx 74.667 \\
\sigma_y &= \sqrt{V_y} = 8.64
\end{aligned}$$

The result for y *is obvious since for reach* i*,* $y_i = 2x_i$*. Alternative way of calculating this:*

$$\begin{aligned}
\overline{x^2} &= (2^2 + 4^2 + 12^2)/3 = \frac{164}{3} = 54\frac{2}{3} \approx 54.6667 \\
V_x &= \overline{x^2} - \overline{x}^2 = \frac{164}{3} - 6^2 = 56/3 \approx 18.667 \\
\sigma_x &= \sqrt{V_x} = 4.32 \\
\overline{y^2} &= (4^2 + 8^2 + 24^2)/3 = \frac{656}{3} \approx 218.667 \\
V_y &= \overline{y^2} - \overline{y}^2 = \frac{656}{3} - 12^2 = \frac{224}{3} \approx 74.667 \\
\sigma_y &= \sqrt{V_y} = 8.64
\end{aligned}$$

- *the covariance.. remembering that:*

$$\begin{aligned}
\mathrm{cov}(x,y) &= \frac{1}{N}\sum(x_i - \overline{x})(y_i - \overline{y}) \\
&= \overline{xy} - \overline{x}\,\overline{y}
\end{aligned}$$

we have again two ways of calculating this

$$\begin{aligned}
\mathrm{cov}(x,y) &= \left[(2-6)(4-12) + (4-6)(8-12) + (12-6)(24-12)\right]/3 \\
&= \frac{112}{3} \approx 37.333
\end{aligned}$$

or

$$\begin{aligned}
\overline{xy} &= (2\cdot 4 + 4\cdot 8 + 12\cdot 24)/3 \\
&= \frac{328}{3} \approx 109.333 \\
\mathrm{cov}(x,y) &= \overline{xy} - \overline{xy} \\
&= \frac{328}{3} - \frac{18}{3}\cdot 12 \\
&= \frac{112}{3} \approx 37.333
\end{aligned}$$

- *And now the correlation coefficient*

$$\begin{aligned}
\rho_{xy} &= \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} \\
&= \frac{112}{3} \cdot \frac{1}{\sqrt{\frac{56}{3}\frac{224}{3}}} = \frac{112}{3} \cdot \frac{\sqrt{3^2}}{\sqrt{56\cdot 224}} \\
&= \frac{112}{\sqrt{56\cdot 224}} = \frac{112}{\sqrt{12544}} = \frac{112}{112} \\
\rho_{xy} &= 1
\end{aligned}$$

The correlation coefficient is exactly +1. This will always be the case when y is proportional to x with a positive constant of proportionality, i.e. $y_i = ax_i \quad \forall i$, with $a > 0$. What will ρ_{xy} be for a negative a?

b)

x	4	1	4
y	2	4	6

Calculated exactly as above. Get

$$\begin{aligned}
\overline{x} &= 3 \\
\overline{y} &= 4 \\
\sigma_x &= \sqrt{\frac{(4-3)^2+(1-3)^2+(4-3)^2}{3}} = \sqrt{2} \approx 1.41 \\
\sigma_y &= \sqrt{\frac{(2-4)^2+(4-4)^2+(6-4)^2}{3}} = \sqrt{\frac{8}{3}} \approx 1.63 \\
\mathrm{cov}(xy) &= \frac{(4-3)(2-4)+(1-3)(4-4)+(4-3)(6-4)}{3} = 0 \\
\rho_{xy} &= 0
\end{aligned}$$

1.2.4 Dice

(a) Suppose you throw two dice, what is the probability that the dice add up to 4 *given* that the first shows a 2?
If the first die shows a 2, then for the dice to add up to 4 the second die must also show a 2. So the probability that the dice add up to 4 given that the first shows a 2 is 1/6. (Q& A from http://philosophy.hku.hk/think/stat)

(b) Suppose you throw two dice, what is the probability that the dice add up to 4 *and* that the first shows a 2?
This is the probability that the first die shows a 2 (which is $1/6$), multiplied by the probability that the dice add up to 4 given *that the first die shows a 2, which we calculated above as $1/6$. Hence the answer is $1/6 \cdot 1/6 = 1/36$. (Q& A from http://philosophy.hku.hk/think/stat)*

1.2.5 Outbreak

Classic question using Bayes theorem A new strain of techno-prions slowly turns humans into tamagotchi toys over a time period of 20 years. The first 10 years usually go un-noticed. The University of Bristol's Medical School has been quick to develop a test to identify those who carry this horrific disease. The test identifies correctly 100% of those who carry the prions, and misidentifies 0.2% of uninfected people as carriers. Assume that 0.01% of the population carry the prions.

a) What is the probability that somebody identified as a carrier by the test does in fact carry the prion?
Notation: a=alarm, c=carrier, h=healthy.
We know:

$$\begin{aligned}
P(a|c) &= 1 \\
P(a|h) &= 2 \cdot 10^{-3} \\
P(c) &= 10^{-4}
\end{aligned}$$

We want $P(c|a)$. Start with Bayes:

$$P(c|a) = P(a|c) \cdot \frac{P(c)}{P(a)}$$

We know P(c), it $P(c) = 0.01\%$. Let's get P(a):

$$\begin{aligned} P(a) &= P(a\&c) + P(a\&h) \\ &= P(a|c)P(c) + P(a|h)P(h) \\ &= P(a|c)P(c) + P(a|h)(1-P(c)) \end{aligned}$$

Now get $P(h|a)$:

$$\begin{aligned} P(c|a) &= P(a|c) \cdot \frac{P(c)}{P(a|c)P(c) + P(a|h)(1-P(c))} \\ &= 1 \cdot \frac{10^{-4}}{1 \cdot 10^{-4} + 2 \cdot 10^{-3} \cdot (1 - 10^{-4})} \\ &= 4.76\%. \end{aligned}$$

So, not very likely.

b) How much does the test need to be improved so that a positive test implies a probability of carrying the prions of at least 95%?
Can't improve $P(a|c)$, so improve $P(a|h)$. Take

$$\begin{aligned} P(c|a) &= P(a|c) \cdot \frac{P(c)}{P(a|c)P(c) + P(a|h)(1-P(c))} \\ & \quad \text{putting in } P(a|c) = 1: \\ &= \frac{P(c)}{P(c) + P(a|h)(1-P(c))} \end{aligned}$$

from prev question and solve for $P(a|h)$ with $P(c|a) = 0.95$:

$$\begin{aligned} P(c|a) &= \frac{P(c)}{P(c) + P(a|h)(1-P(c))} \\ P(c|a)P(c) + P(a|h)P(c|a)(1-P(c)) &= P(c) \\ P(a|h) &= \frac{P(c)(1-P(c|a))}{P(c|a)(1-P(c))} \\ P(a|h) &= \frac{10^{-4} \cdot 0.05}{0.95 \cdot (1-10^{-4})} \\ &= 5.26 \cdot 10^{-6} \end{aligned}$$

So, the mis-id fraction must by reduced by a factor of $2 \cdot 10^{-3} / (5.26 \cdot 10^{-6}) = 380$.

1.2.6 Doping

Victoria plays a sport in which 1% of participants take performance enhancing drugs. She is randomly called for a drug test. The drug test identifies correctly 98% of real drug users, and mis-identifies 3% of non-users as drug users.

a) Victoria tests positive. What is the probability that she actually took drugs?
Notation: a=alarm, d=druggy, c= clean. We want $P(d|a)$.
Start with Bayes:

$$P(d|a) = P(a|d) \cdot \frac{P(d)}{P(a)}$$

We know P(d), it's $P(d) = 1\%$. Let's get P(a):

$$\begin{aligned} P(a) &= P(a\&d) + P(a\&c) \\ &= P(a|d)P(d) + P(a|c)P(c) \\ &= 0.98 \cdot 0.01 + 0.03 \cdot 99 \\ &= 0.0395 \end{aligned}$$

Now get $P(d|a)$:

$$\begin{aligned} P(d|a) &= P(a|d) \cdot \frac{P(d)}{P(a)} \\ &= 0.98 \cdot \frac{0.01}{0.0395} \\ &= 24.8\% \end{aligned}$$

So, not very likely.

b) After the first (positive) test, she is tested immediately again. She tests positive again. What is the probability that she actually took drugs, now (assuming that the probability that the 2nd test returns a correct result is independent of the first test)?
Approach 1 *Same calculation, only now replace $P(d)$ with $P(d_{\text{after 1st test}}) = 24.8\%$ and $P(c)$ with $P(c_{\text{after 1st test}}) = 1 - P(d_{\text{after 1st test}}) = 75.2\%$*

$$P(d_{\text{after 1st test}}|a) = P(a|d_{\text{after 1st test}}) \cdot \frac{P(d_{\text{after 1st test}})}{P(a)}$$

Let's get $P(a)$:

$$\begin{aligned} P(a) &= P(a\&d_{\text{after 1st test}}) + P(a\&c_{\text{after 1st test}}) \\ &= P(a|d_{\text{after 1st test}})P(d_{\text{after 1st test}}) + P(a|c_{\text{after 1st test}})P(c_{\text{after 1st test}}) \\ &= 0.98 \cdot 0.248 + 0.03 \cdot 0.752 \\ &= 0.2656 \end{aligned}$$

Now get $P(d_{\text{after 1st test}}|a)$:

$$\begin{aligned} P(d_{\text{after 1st test}}|a) &= P(a|d_{\text{after 1st test}}) \cdot \frac{P(d_{\text{after 1st test}})}{P(a)} \\ &= 0.98 \cdot \frac{0.248}{0.2656} \\ &= 91.5\% \end{aligned}$$

Quite likely. Still not exactly bullet-proof.
Alternative approach to get the same result:
Starting from the original population where $P(d) = 1\%$, and using that the probability of two positive tests (denoted a_2) for a clean person is $p(a_2|c) = (0.03)^2 = 9 \cdot 10^{-4}$, and for a user it is $p(a_2|d) = 0.98^2 = 0.96$.

$$\begin{aligned} P(d|a_2) &= P(a_2|d) \cdot \frac{P(d)}{P(a_2)} \\ &= P(a_2|d) \cdot \frac{P(d)}{P(a_2|d)P(d) + P(a_2|c)P(c)} \\ &= \frac{0.96 \cdot 0.01}{0.96 \cdot 0.01 + 9 \cdot 10^{-4} \cdot 0.99} \\ &= 91.5\% \end{aligned}$$

which is the same as above.

Chapter 2

Practice Problems Week 2

2.1 Theoretical Distributions

2.1.1 Coins

Level: easy Recall the distribution of obtaining N_H heads when flipping a coin 4 times? If not, look it up in the lecture notes.

$$
\begin{aligned}
P(N_H) &= \left(\frac{1}{2}\right)^4 \binom{4}{N_H} \\
P(0) &= 1/16 \\
P(1) &= 1/4 \\
P(2) &= 3/8 \\
P(3) &= 1/4 \\
P(4) &= 1/16
\end{aligned}
$$

(a) We call the mean of a predicted or theoretical distribution the "expectation value". What is the expectation value $\langle N_H \rangle$ and the standard deviation σ_{N_H}, of this distribution? Mean:

$$
\begin{aligned}
\langle N_H \rangle &= \sum P(N_{H\,i}) N_{H\,i} \\
&= \frac{1}{1} \sum P(N_{H\,i}) N_{H\,i} \\
&= \frac{1}{1}\left(0 \cdot \frac{1}{16} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16}\right) \\
&= 2
\end{aligned}
$$

Variance:

$$
\begin{aligned}
V &= \sum P(N_{H\,i}) \left(N_{H\,i} - \langle N_H \rangle\right)^2 \\
&= \frac{1}{1} \sum P(N_{H\,i}) \left(N_{H\,i} - \langle N_H \rangle\right)^2 \\
&= (0-2)^2 \cdot \frac{1}{16} + (1-2)^2 \cdot \frac{1}{4} + (2-2)^2 \cdot \frac{3}{8} + (3-2)^2 \cdot \frac{1}{4} + (4-2)^2 \cdot \frac{1}{16} \\
&= 1
\end{aligned}
$$

Standard deviation $\sigma = \sqrt{V} = 1$. Also possible to use

$$V = \langle N_H^2 \rangle - \langle N_H \rangle^2$$

Try it!

2.1.2 PDFs

Even PDF's

Level: easy-medium Show that the expectation value of any even PDF is 0. An even function is one that has $f(x) = f(-x)$.

$$\begin{aligned}
\langle x \rangle &= \int_{-\infty}^{\infty} x\, f(x) dx \\
&= \int_{-\infty}^{0} x\, f(x) dx + \int_{0}^{\infty} x\, f(x) dx \\
& \textit{in first integral, } u = -x \\
&= -\int_{\infty}^{0} (-u)\, f(-u) du + \int_{0}^{\infty} x\, f(x) dx \\
&= \int_{\infty}^{0} u\, f(-u) du + \int_{0}^{\infty} x\, f(x) dx \\
&= \int_{\infty}^{0} u\, f(u) du + \int_{0}^{\infty} x\, f(x) dx \\
&= -\int_{0}^{\infty} u\, f(u) du + \int_{0}^{\infty} x\, f(x) dx \\
&= -\int_{0}^{\infty} x\, f(x) dx + \int_{0}^{\infty} x\, f(x) dx \\
&= 0 \quad QED
\end{aligned}$$

Flat

Level: medium Show that the standard diviation of a flat distribution is

$$\sigma_{\text{flat}} = \frac{\text{width}}{\sqrt{12}}$$

Hints: Your distribution is:

$$F(x) = \left\{ \begin{array}{lll} 1/(b-a) & \text{if} & x \in [a, b] \\ 0 & & \text{else} \end{array} \right\}$$

with width $= (b - a)$. Calculate $V_x = \langle x^2 \rangle - \langle x \rangle^2$. It's obvious what $\langle x \rangle$ should be, but do check that you can calculate it.

$$
\begin{aligned}
\langle x \rangle &= \int x F(x) dx \\
&= \int_a^b x/(b-a) dx \\
&= \left[\frac{1}{2}x^2\right]_a^b /(b-a) \\
&= \frac{1}{2}\frac{b^2-a^2}{b-a} \\
&= \frac{1}{2}\frac{(b-a)(b+a)}{b-a} \\
&= \frac{1}{2}(a+b)
\end{aligned}
$$

OK, that was expected. Now $\langle x^2 \rangle$

$$
\begin{aligned}
\langle x^2 \rangle &= \int x^2 F(x) dx \\
&= \int_a^b x^2/(b-a) dx \\
&= \left[\frac{1}{3}x^3\right]_a^b /(b-a) \\
&= \frac{1}{3}\frac{b^3-a^3}{b-a} \\
&= \frac{1}{3}\frac{(b-a)^3-3(ba^2-b^2a)}{b-a} \\
&= \frac{1}{3}\frac{(b-a)^3-3ba(a-b)}{b-a} \\
&= \frac{1}{3}\left((b-a)^2+3ab\right)
\end{aligned}
$$

and finally V

$$
\begin{aligned}
\langle x^2 \rangle - \langle x \rangle^2 &= \frac{1}{3}(b-a)^2+ab-\frac{1}{4}(a+b)^2 \\
&= \frac{1}{3}(b-a)^2+ab-\frac{1}{4}(a+b)^2 \\
&= \frac{1}{3}(a^2+b^2)-\frac{2}{3}ab+ab-\frac{1}{4}(a^2+b^2)-\frac{1}{2}ab \\
&= \frac{1}{12}(a^2+b^2)-\frac{1}{6}ab \\
&= \frac{1}{12}\left(a-b\right)^2 \qquad QED
\end{aligned}
$$

Decay

Level: medium The probability density for an unstable nucleus that exits at time $t_0 = 0$ to decay at a later time t is given by

$$p(t) = \frac{1}{\tau}e^{-t/\tau}$$

(a) Show that this distribution is properly normalised. Keep in mind that the particle cannot have decayed in the past, so $p(t) = 0$ for $t < 0$.

What needs to be shown is:

$$\int_{-\infty}^{+\infty} p(t)\, dt = 1$$

Keeping in mind that $p(t) = 0$ for $t < 0$ this is:

$$\int_{0}^{+\infty} \frac{1}{\tau} e^{-t/\tau}\, dt = 1$$

Calculating the integral:

$$\begin{aligned}
\int_{0}^{+\infty} \frac{1}{\tau} e^{-t/\tau}\, dt &= \frac{1}{\tau}\left[-\frac{1}{\tau} e^{-t/\tau}\right]_0^{\infty} \\
&= -\left(e^{-\infty} - e^{0}\right) \\
&= -(0-1) \\
&= 1 \quad \text{QED}
\end{aligned}$$

(b) The cumulative probability distribution F for a pdf p is defined as

$$F(t) \equiv \int_{-\infty}^{t} p(t')\, dt'$$

What is the cumulative probability distribution for $p(t) = \frac{1}{\tau} e^{-t/\tau}$

Same calcualation as above, except that you replace the upper limit with t (and make sure we use different variables for limit and integration variable, so we re-name the latter as t').

$$\begin{aligned}
F(t) &\equiv \int_{-\infty}^{t} p(t')\, dt' \\
&= \int_{0}^{t} \frac{1}{\tau} e^{-t'/\tau}\, dt' \\
&= \frac{1}{\tau}\left[-\tau e^{-t'/\tau}\right]_0^{t} \\
&= -\left(e^{-t/\tau} - e^{0}\right) \\
&= 1 - e^{-t/\tau}
\end{aligned}$$

(c) What is the expectation value and the standard deviation of $p(t)$?

- *Expectation value:*

$$\langle t \rangle = \int_{0}^{\infty} t \frac{1}{\tau} e^{-t/\tau}\, dt$$

There are now several ways of solving this. First, using integration by parts:

$$\begin{aligned}
\langle t \rangle &= \frac{1}{\tau}\int_0^\infty t e^{-t/\tau}\,dt \\
& \quad \textit{integration by parts, } u = t, v' = e^{-t/\tau} \\
&= \frac{1}{\tau}\left(\left[t(-\tau)e^{-t/\tau}\right]_0^\infty - \int_0^\infty (-\tau)e^{-t/\tau}\,dt\right) \\
&= \frac{1}{\tau}\left(0 + \tau\left[(-\tau)e^{-t/\tau}\right]_0^\infty\right) \\
&= \frac{1}{\tau}\left(-\tau^2\,(0-1)\right) \\
&= \tau
\end{aligned}$$

Alternative way: Define $\frac{-1}{\tau} = a$. *Then*

$$\begin{aligned}
\langle t \rangle &= \frac{1}{\tau}\int_0^\infty t e^{-t/\tau}\,dt \\
&= \frac{1}{\tau}\int_0^\infty t e^{at}\,dt \\
&= \frac{1}{\tau}\frac{d}{da}\int_0^\infty e^{at}\,dt \\
&= \frac{1}{\tau}\frac{d}{da}\left[\frac{1}{a}e^{at}\right]_0^\infty \\
&= \frac{1}{\tau}\frac{d}{da}\frac{1}{a}(0-1) \\
&= -\frac{1}{\tau}\frac{d}{da}\frac{1}{a} \\
&= \frac{1}{\tau}\frac{1}{a^2} \\
&= \frac{1}{\tau}\tau^2 \\
&= \tau
\end{aligned}$$

- *Expectation value of* t^2, *i.e.* $\langle t^2 \rangle$

$$\langle t \rangle = \int_0^{t_0} t^2 \frac{1}{\tau} e^{-t/\tau}\,dt$$

Using partial integration:

$$\begin{aligned}
\langle t\rangle &= \frac{1}{\tau}\int_0^\infty t^2 e^{-t/\tau}\,dt \\
&= \frac{1}{\tau}\left(\left[t^2(-\tau)e^{-t\tau}\right]_0^\infty - \frac{1}{\tau}\int_0^\infty (-\tau)2te^{-t/\tau}\,dt\right) \\
&= 0 + 2\tau\frac{1}{\tau}\int_0^\infty te^{-t/\tau}\,dt \\
& \quad \textit{using prev. result for integral:} \\
&= 2\tau\tau = 2\tau^2
\end{aligned}$$

Alternative way: Define $\frac{-1}{\tau} = a$. Then

$$\begin{aligned}
\langle t\rangle &= \frac{1}{\tau}\int_0^\infty t^2 e^{-t/\tau}\,dt \\
&= \frac{1}{\tau}\int_0^\infty t^2 e^{at}\,dt \\
&= -\frac{1}{\tau}\frac{d^2}{da^2}\frac{1}{a} \\
&= \frac{1}{\tau}\frac{d}{da}\frac{1}{a^2} \\
&= \frac{1}{\tau}\frac{d}{da}\frac{-2}{a^3} \\
&= \frac{1}{\tau}(-2)(-\tau)^3 \\
&= 2\tau^2
\end{aligned}$$

- *So the variance is $V = \langle t^2\rangle - \langle t\rangle^2 = \tau^2$ and the standard deviation $\sigma = \sqrt{V} = \tau$.*

(d) Find the co-ordinate transformation that transforms a flat distribution between 0 and 1 (PDF in question 2.1.2 with $a = 0$, $b = 1$) to the above exponential. If we call the flatly-distributed random variable x, you want to find $t(x)$ such that

$$p_{\text{flat}}(t)\,dt = p(x)\,dx$$

where $p(t) = \frac{1}{\tau}e^{-t/\tau}$ and

$$p_{\text{flat}}(x) = \left\{\begin{array}{ll} 1 & \text{if} \quad x \in [0,1) \\ 0 & \text{else} \end{array}\right\}$$

The way to solve this is to first find the inverse transformation, $x(t)$, and then solve this for t to find what you really want, $t(x)$. The first step is to put in $p(t)$ and $p_{\text{flat}}(x)$ and re-arrange things a bit.

$$\frac{dx}{dt} = \frac{1}{\tau}e^{-t/\tau}$$

You should be able to take it from here. This will give you $x(t)$. Now solve for t to get $t(x)$. Use your integration constants wisely - so that you map a distribution between 0 and 1 to a distribution between 0 and ∞.

$$\begin{aligned}
\frac{dx}{dt} &= \frac{1}{\tau}e^{-t/\tau} \\
x(t) &= -e^{-t/\tau} + C \\
C - x &= e^{-t/\tau} \\
\log(C - x)) &= -\frac{t}{\tau} \\
t(x) &= -\tau \log(C - x)
\end{aligned}$$

This is my co-ordinate transformation, except that I still have the undetermined integration constant C in there. This is determined by the fact that I want t to be between 0 and ∞, therefore $C - x$ must be between 0 and 1. Given that x varies between 0 and 1, for $C - x \in [0, 1]$ we require $C = 1$. So finally we get:

$$t(x) = -\tau \log(1 - x)$$

You can try it out using a random number generator for your flatly-distributed x, just calculate $-\tau \log(1 - x)$ for some value of τ and see how this is distributed. Our download the jupyter notebook for this question, which you can find here.

Gaussian

Level: medium to hard The Gaussian PDF is given by:

$$g(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This notation means that $g(x; \mu, \sigma)$ is a function of x, with the parameters μ, σ. You could also simply write $g(x)$.

You can assume that this is properly normalised, i.e. $\int\limits_{-\infty}^{+\infty} g(x)\ dx = 1$ (you might or might not need this information).

(a) Show that all integrals over $g(x; \mu, \sigma)$ can be expressed in terms of the gaussian with $\mu = 0$ and $\sigma = 1$:

$$\int\limits_{a}^{b} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int\limits_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\ du$$

simple co-ordinate transformation $u = \frac{x-\mu}{\sigma}$.

(b) Show that its mean is μ. Hint: Use the result from the previous question and an appropriate co-ordinate transformation.

$$\int\limits_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int\limits_{-\infty}^{\infty} (\sigma \cdot u + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\ du$$

where we used $u = \frac{x-\mu}{\sigma}$ with $du = \frac{dx}{\sigma}$. Now we use that the integral from $-\infty$ to ∞ of an odd function (i.e. a function $f(x)$ with $f(-x) = -f(x)$), therefore

$$\int\limits_{-\infty}^{\infty} u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\ du = 0$$

So we are let with

$$\int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \mu \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du = \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du = \mu \cdot 1 = \mu$$

where we used the $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du = 1$.

(c) Advanced question: Show that the standard deviation of the Gaussian is (as the variable name suggests) σ

$V = \langle x^2 \rangle - \langle x \rangle^2$. *We know* $\langle x \rangle$. *Now calculate* $\langle x^2 \rangle$*:*

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Now do co-ordinate transformation to make it look easier:

$$u = \frac{x-\mu}{\sigma} \quad dx = \sigma\, du$$

With this:

$$\begin{aligned} \langle x^2 \rangle &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma \cdot u + \mu)^2 e^{-\frac{u^2}{2}} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma^2 u^2 + 2\sigma\mu\, u + \mu^2) e^{-\frac{u^2}{2}}\, du \end{aligned}$$

Term by term:

- $$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 2\sigma\mu\, u e^{-\frac{u^2}{2}}\, du$$

  vanishes - odd function over symmetric interval.

- $$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu^2 e^{-\frac{u^2}{2}}\, du = \mu^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}\, du = \mu^2$$

  due to normalisation.

- *To solve the remaining one, we use a trick. We introduce a dummy parametr* a *and express the integral as the derivative of* a *at* $a = 1$*:*

$$\left.\frac{d}{da}\right|_{a=1} (-2e^{-\frac{au^2}{2}}) = u^2 e^{-\frac{u^2}{2}}$$

Where the $\left.\frac{d}{da}\right|_{a=1}$ *means "first calculate the derivative, and then evaluate it at* $a = 1$.*" With this:*

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}\sigma^2 \int_{-\infty}^{\infty} u^2 e^{-\frac{u^2}{2}}\, du &= -2\frac{1}{\sqrt{2\pi}}\sigma^2 \left.\frac{d}{da}\right|_{a=1} \int_{-\infty}^{\infty} e^{-\frac{au^2}{2}}\, du \\
&\quad \textit{coord trafo: } v = \sqrt{a}u \\
&= -2\frac{1}{\sqrt{2\pi}}\sigma^2 \left.\frac{d}{da}\right|_{a=1} \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-\frac{v^2}{2}}\, dv \\
&= -2\sigma^2 \left.\frac{d}{da}\right|_{a=1} \frac{1}{\sqrt{a}} 1 \\
&= -2\sigma^2 (-\frac{1}{2}) \\
&= \sigma^2
\end{aligned}$$

So we have

$$\langle x^2 \rangle = \sigma^2 + \mu^2$$

and consequently

$$V = \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

and $\sigma = \sqrt{V} = \sigma$, *as expected.*

Convolution

Level: fairly tricky - but you will meet convolutions again in part 2 of the course, so why not familiarise yourself with them, now. Consider random variable $z = x + y$, where x is distributed according to the PDF p_x and y is distributed according to the PDF p_y. The PDF of z is given by:

$$p_z(z) = \int_{-\infty}^{\infty} p_x(x) p_y(z - x) dx = \int_{-\infty}^{\infty} p_y(x) p_x(z - x) dx$$

If x and y are both distributed according to a flat distribution between 0 and 1 (see question 2.1.2), what is p_z?

$$\begin{aligned}
p_z(z) &= \int_{-\infty}^{\infty} p_{\text{flat}}(x) p_{\text{flat}}(z - x) dx \\
&= \int_{0}^{1} p_{\text{flat}}(z - x) dx
\end{aligned}$$

- *For* $z \leq 0$ $p(z) = 0$
- *for* $z \geq 2$ $p(z) = 0$

- *Else: u=z-x x=z-u*

$$\begin{aligned} p_z(z) &= \int_0^1 p_{\text{flat}}(z-x)dx \\ &= -\int_z^{z-1} p_{\text{flat}}(u)du \\ &= \int_{z-1}^{z} p_{\text{flat}}(u)du \\ &= \int_{\max(z-1,0)}^{\min(z,1)} p_{\text{flat}}(u)du \\ &= \int_{\max(z-1,0)}^{\min(z,1)} 1\ du \end{aligned}$$

The result depends on the explicit form of the integration limits, which are $u_{lower} = \max(z-1,0)$ *and* $u_{upper} = \min(z,1)$.

- *For* $z < 0$, *the integration limits are* $u_{lower} = u_{upper} = 0$ *and hence* $p_z(z) = 0$
- *for* $z > 2$, *the integration limits are* $u_{lower} = u_{upper} = 1$ *and hence also* $p_z(z) = 0$
- *For* $0 < z < 1$: $u_{lower} = 0, u_{upper} = z$ *and therefore* $p_z(z) = z$
- *For* $1 < z < 2$: $u_{lower} = (z-1), u_{upper} = 1$ *and therefore* $p_z(z) = 2 - z$.

So the solution is a triangle that peaks at $p(z=1) = 1$. *You can try it out using this jupyter notebook.*

2.1.3 Calculating Probabilities

Life insurance

Level: Easy-medium once you've done the Binomial distribution The sort of thing that goes through the mind of an insurance agent when he asks you how old you are: If the probability that any person thirty years old will be dead within 10 years is $p = 0.01$, find the probability that out of a group of 10 such people
Clearly a case for the binomial distribution.

$$P(r) = \left(\begin{array}{c} n \\ r \end{array} \right) p^r (1-p)^{n-r} = \frac{n!}{(n-r)!\, r!} p^r (1-p)^{n-r} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{r!} p^r (1-p)^{n-r}$$

with $n = 10$.
You'll find the the Poisson with $\lambda = n \cdot p = 0.1$ *will give you a decent approximation.*

a) none

$$P(0) = (1-p)^{10} = 90.44\%$$

b) exactly one

$$P(1) = n \cdot p(1-p)^{n-1} = 9.04\%$$

c) not more than one

$$P(0) + P(1) = 99.48\%$$

d) more than one

$$1 - (P(0) + P(1)) = 0.52\%$$

e) at least one

$$1 - P(0) = 9.56\%$$

will be dead within 10 years.

Parity Bit

Level: medium (once you've done the binomial distribution Let's say that the probability that a single bit (which can have the values "I" or "0") is transmitted wrongly (i.e. what is sent as "I" arrives as "0" and vice versa) over the internet is $p = 10^{-6}$ (completely made-up number).

a) What is the probability that an ASCII character of 7 bits will have at least one wrong bit?
   Binomial:

$$P(r) = \binom{n}{r} p^r (1-p)^{n-r} = \frac{n!}{(n-r)!\,r!} p^r (1-p)^{n-r} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{r!} p^r (1-p)^{n-r}$$

   with $n = 7$, $p = 10^{-6}$

$$P(\textit{at least 1}) = 1 - P(0) = 1 - (1-p)^7 = 7.00 \cdot 10^{-6}$$

   If your calculator gave you 0, *it hasn't got enough significant digits for this calculation. Then use*

$$P(\textit{at least 1}) = 1 - P(0) = 1.0 - (1-p)^7 = 1 - (1 - 7p + \mathcal{O}(p^2)) \approx 7p = 7.00 \cdot 10^{-6}$$

   where we've used that $p \ll 1$

b) How many characters do you expect to be wrong in a $1\mathrm{MB} = 2^{20}$ bytes of text?

   - *Single byte:* $p_{\mathrm{byte}} = 7 \cdot 10^{-6}$
   - 2^{20} *bytes:*

$$\langle N_{\mathrm{wrong}} \rangle = 2^{20} \cdot 7 \cdot 10^{-6} = 7.3$$

c) One way to reduce the error rate in such transmissions is the so-called parity bit. The sender adds to each 7-bit ASCII character one more bit. This is done such that to over-all number of bits set to "I" is even. So the 7-bit character "0III000" would have a partity bit of "I" added, giving "0III000I":

   0III000
   →0III000I ,

   while "0II0000" would have a "0" added, giving "0II00000":

   0II0000
   →0II00000 .

The recipient will check the number of "I" bits and request re-transmission if it finds an odd number. This method will find single-bit failures. How many wrong characters to you expect in a 1MB text, now?
Now we have again a binomial distribution, however instead of $n = 7$ we have $n = 8$. This method can detect single bit failures (and any odd number of bit failures), but not 2-bit failures (or any even number of bit failures). We can safely ingore any larger contribution from higher-order terms such as three or four bit failures since they are supressed relative to the "leading order" term of 2-bit failures by a factor of $p = 7 \cdot 10^{-6}$, so they are a correction somewhere in the 5th significant digit of our result. (For the same reason we don't have to worry about the details of re-transmission, we can safely assume that a character once identified as wrongly transmitted, is correctly transmitted the next time - the correction due to this relative to the number we will calculate is again less than 1/100,000, so somewhere in the 5th significant digit)

- *So, while actually we need calculate*

$$P(\text{2, 4, 6, 8 out of 8 bits wrong; for n=8})$$

in this special case, as justified above, it is actually a sufficiently good approximation to calculate only

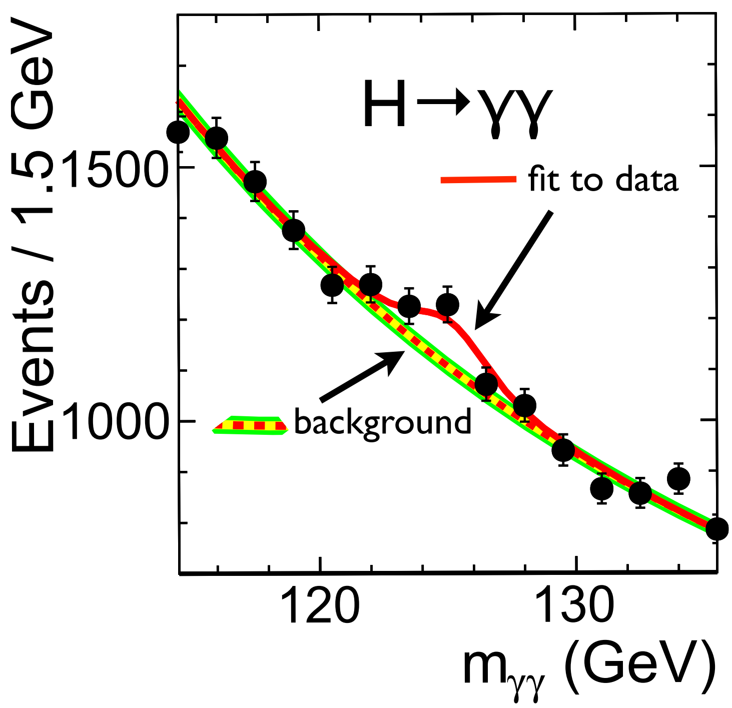
$$P(\text{2 bits wrong; n=8}) = \binom{8}{2} p^2(1-p)^6 = 2.80 \cdot 10^{-11}$$

- *So now the expected number of wrong bytes in a 2^{20} byte text is:*

$$\langle N_{\text{wrong}} \rangle = 2.9 \cdot 10^{-5}$$

So it's a huge effect - it turns something quite useless (any major text - and any long computer programme - would have errors in it) to something that's quite OK. I suspect that even the error rate we calculated with parity bit is still unacceptably high and the actual error rate is much lower.

2.1.4 Confidence Levels: Higgs Search



The plot above shows the result of the CMS experiment's search for the decay of a Higgs meson to two photons. It shows the number of events found in bins of the reconstructed Higgs mass, given in GeV. Apart from the data themselves, the plot also shows a fit to the data (which you can ignore) and the expected background distribution. Around $125\,\mathrm{GeV}$, you can see an excess of events compared to the expected background distribution.

Estimate the probability that the excess above the exponential background distribution is a statistical fluctuation. Please ignore the fit and the error bars. Note that, near $125\,\mathrm{GeV}$, there is only one bin which significantly deviates from the background curve. Assume this bin has 1258 events, while we expect 1156 events from background processes.

a) What is the probability to find such a fluctuation in a given bin?

b) Now calculate the probability of seeing such a fluctuation given that the researchers searched in 100 bins.

This is a $\frac{1258-1156}{\sqrt{1156}} = 3\sigma$ *effect, the probability of such an excess in a given bin is* $1.3 \cdot 10^{-3}$ *(formula table, 1-sided Gaussian). The probability to find this at least once in 100 bins is* $1-(1-1.3\cdot 10^{-3})^{100} = 12\%$.

2.2 Supplementary Statistics Questions

These are for additional practice

2.2.1 Combinatorics

1. The Zimbabwean Cabinet has 30 ministries to be divided between three parties: Zanu-PF (14), Main MDC (13), and MDC (Mutambara) (3).

   i) How many ways are there to allocate the various ministries to the different parties, given these constraints?

   How many ways are there to put 30 things (ministries) into 3 boxes (parties) each with $n_1 = 14$, $n_2 = 13$, $n_3 = 3$ *places?*

   $$N = \frac{n!}{n_1!\,n_2!\,n_3!} = \frac{30!}{14!\cdot 13!\cdot 3!} = 8.14\cdot 10^{10}$$

   ii) If neither Robert Mugabe's Zanu-PF nor the opposition (Main MDC and MDC (Mutambara)) are allowed to have both, defence and home affairs (which includes police), how many ways are there now to divide the ministries between the parties?

   - *Number of ways in which Mugabe's Zanu PF can have both:*

   $$N_{RM} = \frac{n!}{n_1!\,n_2!\,n_3!} = \frac{28!}{12!\cdot 13!\cdot 3!} = 1.7\cdot 10^{10}$$

   - *Number of ways in which the Main MDC can have both:*

   $$N_{MDC-main} = \frac{n!}{n_1!\,n_2!\,n_3!} = \frac{28!}{14!\cdot 11!\cdot 3!} = 1.5\cdot 10^{10}$$

   - *Number of ways in which the MDC-Mutaba can have both:*

   $$N_{MDC-mutaba} = \frac{n!}{n_1!\,n_2!\,n_3!} = \frac{28!}{14!\cdot 13!\cdot 1!} = 5.6\cdot 10^{8}$$

   - *Number of ways in which the Main MDC and MDC-Mutaba can each have exactly one*

   $$N_{MDC-shared} = 2\cdot\frac{n!}{n_1!\,n_2!\,n_3!} = \frac{28!}{14!\cdot 12!\cdot 2!} = 3.7\cdot 10^{9}$$

   - *So, total:*

   $$N_{\mathrm{constraint}} = N - N_{RM} - N_{MDC-main} - N_{MDC-mutaba} - N_{MDC-shared} = 4.6\cdot 10^{10}$$

2.2.2 Molecules*

Level: medium-hard

1) According to the equipartition theorem, the average kinetic Energy of gas molecules at temperature T is $\overline{E_{kin}} =$ ____________. (Look up the equipartition theorem if necessary and don't forget that the speed of the particle corresponds to 3 degrees of freedom).
$\frac{3}{2}k_B T$ What is the mean speed squared, $\overline{|v|^2}$, for a given molecular mass m?
$\frac{3kT}{m}$

2) From Maxwell's speed distribution, we know that the mean speed of a molecule at temperature T is $\overline{|v|} = \sqrt{\frac{8k_B T}{\pi m}}$. Compare this to the previous result. Why is it OK to calculate the $\overline{|v|^2}$ from the kinetic energy $\overline{E_{kin}}$, while it is not OK to calculate $\overline{|v|}$ from it?
Because the mean is "linear", i.e. it behaves like this:

$$\overline{\lambda x + \mu y} = \lambda \overline{x} + \mu \overline{y}$$

where λ and μ are fixed numbers, and x and y represent data from two different data sets. This of course also implies that

$$\overline{\lambda x} = \lambda \overline{x}.$$

There is a linear relationship between E and v^2 (which is $E = \frac{1}{2}m(v^2)$, i.e. in the formalism above, $\lambda = \frac{1}{2}m$ and $x = v^2$). Therefore

$$\overline{E} = \frac{1}{2}m\overline{v^2}$$

However, the relationship between v and v^2 is $v = \sqrt{v^2}$, which is not linear. Hence, in general $\overline{v} \neq \sqrt{\overline{v^2}}$ (of course one cannot exclude the possibility that by pure chance the two numbers turn out to be the same).

3) Show that if, and only if the fluctuations about the mean $|v|^2$ are small, then $\overline{|v|} \approx \sqrt{\overline{|v|^2}}$. Instructions:

a) Obviously, $v = \sqrt{v^2}$. Express v^2 as the $\overline{|v|^2} + \Delta v^2$. Perform a Taylor expansion of $v = \sqrt{|v|^2}$ around $|v|^2 = \overline{|v|^2}$ up to second order (i.e. get sth that looks like $a + b\Delta v^2 + c\left(\Delta v^2\right)^2$)
Taylor expansion:

$$f(x_0 + \Delta x) = f(x_0) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=x_0} \cdot \Delta x + \frac{1}{2}\left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x=x_0} \cdot (\Delta x)^2 + \dots$$

Here we have $x = v^2$ and $x_0 = \overline{x} = \overline{v^2}$, and $f(x) = \sqrt{x}$, and therefore

$$\begin{aligned} f(x) &= \sqrt{x} \quad \textit{(with } x = v^2\textit{)} \\ \frac{\mathrm{d}f}{\mathrm{d}x} &= \frac{1}{2}\frac{1}{\sqrt{x}} \\ \frac{\mathrm{d}^2 f}{\mathrm{d}x^2} &= -\frac{1}{4}\frac{1}{x^{3/2}} \end{aligned}$$

Plugging it all in:

$$\begin{aligned} v &= \sqrt{x_0} + \frac{1}{2\sqrt{x_0}}\Delta x - \frac{1}{8\left(x_0\right)^{3/2}}\left(\Delta x\right)^2 \\ &= \sqrt{\overline{x}} + \frac{1}{2\sqrt{\overline{x}}}\Delta x - \frac{1}{8\left(\overline{x}\right)^{3/2}}\left(\Delta x\right)^2 \\ &= \sqrt{\overline{|v|^2}} + \frac{1}{2\sqrt{\overline{|v|^2}}}\Delta|v|^2 - \frac{1}{8\left(\overline{|v|^2}\right)^{3/2}}\left(\Delta|v|^2\right)^2 \end{aligned}$$

b) Now take the mean of that. Note that $\overline{\Delta|v|^2}$ is the 1st central moment of $|v|^2$ - what happens to that term? What does the $\overline{(\Delta|v|^2)^2}$ term represent?

$$
\begin{aligned}
\bar{v} &= \overline{\sqrt{\overline{|v|^2}}} + \overline{\frac{1}{\sqrt{\overline{|v|^2}}}\Delta|v|^2} - \overline{\frac{1}{8\left(\overline{|v|^2}\right)^{3/2}}\left(\Delta|v|^2\right)^2} \\
&= \sqrt{\overline{|v|^2}} + \frac{1}{\sqrt{\overline{|v|^2}}}\overline{\Delta|v|^2} - \frac{1}{8\left(\overline{|v|^2}\right)^{3/2}}\overline{\left(\Delta|v|^2\right)^2}
\end{aligned}
$$

Let's look at each term with a $\overline{bar}$ over it in turn:

$$
\begin{aligned}
\overline{\sqrt{\overline{v^2}}} &= \sqrt{\overline{v^2}} \quad \textit{just a constant} \\
\overline{\Delta|v|^2} &= \overline{|v|^2 - \overline{|v|^2}} \\
&= \overline{|v|^2} - \overline{\overline{|v|^2}} \\
&= \overline{|v|^2} - \overline{|v|^2} \\
&= 0 \\
\overline{(\Delta|v|^2)^2} &= \overline{\left(|v|^2 - \overline{|v|^2}\right)^2} \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(|v|^2 - \overline{|v|^2}\right)^2 \\
&= Var(|v|^2) = \sigma^2_{|v|^2}
\end{aligned}
$$

Putting these results back in:

$$
\begin{aligned}
\bar{v} &\approx \sqrt{\overline{|v|^2}} - \frac{1}{8\left(\overline{|v|^2}\right)^{3/2}}\overline{(\Delta|v|^2)^2} \\
&= \sqrt{\overline{|v|^2}} - \frac{1}{8\left(\overline{|v|^2}\right)^{3/2}}\mathrm{Var}\left(|v|^2\right) \\
&= \sqrt{\overline{|v|^2}} - \frac{1}{8\sqrt{\overline{|v|^2}}}\frac{\mathrm{Var}\left(|v|^2\right)}{\overline{|v|^2}}
\end{aligned}
$$

c) Hence, if the Variance of $|v|^2$ is small compared to $\overline{|v|^2}$, and *only* then, is $\overline{|v|} \approx \sqrt{\overline{|v|^2}}$ an acceptable approximation.

2.2.3 CLT

A random variable X is the sum of N independent random variables y_i with $i = 1, \ldots, N$:

$$X = \sum_{i=1}^{N} y_i$$

i) Prove that

$$\langle X \rangle = \sum \langle y_i \rangle$$

(this is easy - use linearity of expectation values)

$$
\begin{aligned}
X &= \sum_i y_i \\
\langle X \rangle &= \langle \sum_i y_i \rangle \\
\langle X \rangle &= \sum_i \langle y_i \rangle
\end{aligned}
$$

ii) Prove that the standard deviation of X is given by

$$\sigma_x = \sqrt{\sum \sigma^2(y_i)}$$

where $\sigma(y_i)$ is the standard deviation of the random variable y_i. Hint: Start with

$$V(X) = \langle (X - \langle X \rangle)^2 \rangle$$

$$
\begin{aligned}
V(X) &= \langle (X - \langle X \rangle)^2 \rangle \\
&= \langle \left(\sum_i y_i - \sum_i \langle y_i \rangle \right)^2 \rangle \\
&= \langle \left(\sum_i (y_i - \langle y_i \rangle) \right)^2 \rangle \\
&= \langle \left(\sum_i (y_i - \langle y_i \rangle) \right) \left(\sum_j (y_j - \langle y_j \rangle) \right) \rangle \\
&= \langle \sum_{ij} (y_i - \langle y_i \rangle) ((y_j - \langle y_j \rangle)) \rangle \\
&= \sum_{ij} \langle \ (y_i - \langle y_i \rangle) ((y_j - \langle y_j \rangle)) \ \rangle \\
&= \sum_{ij} \operatorname{cov}(y_i, \, y_j) \\
&\quad \textit{We stated that all the } y_i \textit{ are independent - so the covariance between} \\
&\quad y_i \textit{ and } y_j \textit{ vanishes, unless of course } i = j\textit{:} \\
&= \sum_i \operatorname{cov}(y_i, y_i) \\
&\quad \textit{The covariance of a parameter with itself is simply the variance} \\
&= \sum_i V_i \\
\sigma_X &= \sqrt{V_x} = \sqrt{\sum_i \sigma^2_{y\,i}} \qquad \textit{QED}
\end{aligned}
$$

Chapter 3

Practice Problems Week 3

3.1 Vector Basics

3.1.1 What animal is it?

We denote vectors with arrows on top, like this: $\vec{A}$. Everything else is just a simple number.

(a) Which of the expressions below is

- a vector
- a scalar (simply a number)
- nonsense?

Remember that the dot: $\cdot$, between vectors is the dot-product. The same symbol between scalars is the usual product between to numbers, and between scalars and vectors it is normal scalar multiplicaton of vectors (which changes the length, but not the direction of vectors). In the latter two cases you can omit the dot - you cannot between vectors, to avoid confusion with the cross product. The cross $\times$ indicates the cross-product and only makes sense between vectors.

(i) $\lambda\vec{A}$
vector

(ii) $\vec{A}\times\vec{B}$
vector

(iii) $\vec{A}\cdot\vec{B}$
scalar

(iv) $\vec{C}\cdot\left(\vec{A}\times\vec{B}\right)$
scalar

(v) $\left(\vec{C}\cdot\vec{A}\right)\times\vec{B}$
nonsense

(vi) $\vec{C}\cdot\vec{A}\times\vec{B}$
scalar (interpet as $\vec{C}\cdot\left(\vec{A}\times\vec{B}\right)$)

(vii) $\left(\vec{C}\cdot\vec{A}\right)\vec{B}$
vector

(viii) $\left(\vec{A}\times\vec{B}\right)\times\vec{C}$
vector

(ix) Why is this nonsense: $\vec{A}\times\vec{B}+\vec{A}\cdot\vec{B}$
Can only add vectors and vectors, as well as scalars and scalars, but never vectors and scalars.

(b) Which of the below is zero (either scalar-zero 0 or vector zero $\vec{0}$), which one is not zero, which one can be either, depending on the value of $\vec{A}$ and $\vec{B}$. You can assume that neither $\vec{A}$ nor $\vec{B}$ themselves are $\vec{0}$:

(i) $\vec{A} \cdot \vec{A}$
non-zero

(ii) $\vec{A} \times \vec{A}$
$\vec{0}$

(iii) $\left(\vec{A} \times \vec{B}\right) \cdot \vec{A}$
It's 0 because $\vec{A} \times \vec{B}$ is perpendicular to $\vec{A}$ and the dot-product between perpendicular vectors is 0. Can also be seen like this: the triple-dot product between three vectors represents the volume spanned by the three. Here we only have two vectors - they always lie in one plane, so the volume is 0.

(iv) $\left(\vec{A} \times \vec{B}\right) \times \vec{A}$
Could go either way. Definetely not 0 if $\vec{A} \times \vec{B} \neq \vec{0}$.

(v) $\lambda\vec{A} + \mu\vec{B}$ for $\mu \neq 0$, $\lambda \neq 0$
Can be either. Definetely non-zero only if $\vec{A}$ and $\vec{B}$ are linearly independent - but we don't know that.

(vi) $\vec{A} \times \vec{B} + \left(\vec{A} \cdot \vec{B}\right)\vec{B}$ Hint: The answer to this one is *not* "can be either".
Cannot be 0 because

- *$\vec{A} \times \vec{B}$ and $\vec{B}$ are linearly independent.*
- *Given that $\vec{A}$ and $\vec{B}$ are each non-zero, $\vec{A} \times \vec{B}$ can only be zero if $\vec{A}$ and $\vec{B}$ are parallel, while $\vec{A} \cdot \vec{B}$ can only be zero if they are perpendicular. Both cannot be the case at the same time.*
- *So we have the sum of two linearly independent vectors, at least one of which is non-zero, so the result is non-zero.*

3.1.2 Vector operations

(a) Calculate $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$

$\begin{pmatrix} -4 \\ 8 \\ -4 \end{pmatrix}$

(b) Draw the vectors $\vec{A} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ and $\vec{B} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$. These "span" a paralellogram (diamond shape). What is its area?

$$\text{area} = \left|\vec{A} \times \vec{B}\right| = \left|\begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix}\right| = 3$$

(c) Calculate $\vec{A} \cdot \vec{B}$ for $\vec{A} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ and $\vec{B} = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}$

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} = 0 \cdot 1 + 1 \cdot 4 + 2 \cdot 0 = 4$$

(d) Calculate the length of $\vec{A}$ and $\vec{B}$ in the previous questions
length = sqare-root of dot-product with itself:

$$\left|\vec{A}\right| = \sqrt{0\cdot 0 + 1\cdot 1 + 2\cot 2} = \sqrt{5} \approx 2.23$$

same for $\vec{B}$:

$$\left|\vec{A}\right| = \sqrt{1\cdot 1 + 4\cdot 4 + 0\cot 0} = \sqrt{17} \approx 4.1$$

(e) Calculate the angle between $\vec{A}$ and $\vec{B}$ from the previous two questions.
Let's call the angle θ. With this:

$$\vec{A}\cdot\vec{B} = \left|\vec{A}\right|\left|\vec{B}\right|\cos\theta$$

Therefore

$$\cos\theta = \frac{4}{\sqrt{5}\sqrt{17}} = 0.43$$

and

$$\theta = 1.1$$

(which is 64^o).

(f) What this the unit vector in the direction of $\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$?
Divide by the vector's length to get the unit vector:

$$\vec{u} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \frac{1}{\left|\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}\right|}$$

where

$$\left|\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}\right| = \sqrt{\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}\cdot\begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}} = \sqrt{25} = 5$$

So

$$\vec{u} = \begin{pmatrix} 3/5 \\ 4/5 \\ 0 \end{pmatrix}$$

(g) Applying a constant force $\vec{F} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} N$ to move, in a straight line an object from point $\vec{A} = \begin{pmatrix} 10 \\ 11 \\ 12 \end{pmatrix} m$ to point $\vec{B} = \begin{pmatrix} 10 \\ 11 \\ 13 \end{pmatrix} m$: How much work is done?
Work = force in direction of path times path length. Usually require integration but the path is straight and the force constant, so we can just multiply things together:

$$W = \vec{F}\cdot\left(\vec{B}-\vec{A}\right) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} N \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} m = 3J$$

3.2 Other Math Basics

(a) Use a Taylor/McLauren expansion to show that

$$\frac{1}{1+x} \approx 1 - x + x^2$$

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n f(x)}{dx^n}\Big|_{x=0} x^n \approx f(0) + \frac{df(x)}{dx}\Big|_{x=0} x + \frac{1}{2}\frac{d^2 f(x)}{dx^2}\Big|_{x=0} x^2$$

with $\frac{df}{dx} = \frac{-1}{(1+x)^2}$ *and with* $\frac{d^2 f}{dx^2} = \frac{2}{(1+x)^3}$

$$f(x) \approx f(0) + (-1) \cdot x \frac{1}{2} \cdot 2 \cdot x^2 = 1 - x + x^2$$

(b) Use a Taylor/McLauren expansion to show that, for values of x near 1,

$$\ln(x) \approx \frac{-3}{2} + 2x - \frac{1}{2}x^2$$

Hint: Expand around $x_0 = 1$.

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n f(x)}{dx^n}\Big|_{x=x_0} (x - x_0)^n \approx f(x_0) + \frac{df(x)}{dx}\Big|_{x=x_0} (x - x_0) + \frac{1}{2}\frac{d^2 f(x)}{dx^2}\Big|_{x=x_0} (x - x_0)^2$$

with $\frac{df}{dx} = \frac{1}{x}$, $\frac{d^2 f}{dx^2} = \frac{-1}{x^2}$,

$$\ln(x) \approx (x-1) - \frac{1}{2}(x-1)^2 = x - 1 - \frac{1}{2}x^2 - \frac{1}{2} + x = \frac{-3}{2} + 2x - \frac{1}{2}x^2$$

(c) Calculate the following integral

$$\iint\limits_{\text{unit disk}} (x^2 + y^2) dx\, dy$$

over the unit disk. Hint: Think carefully about the integration limits, especially if you do it in cartesian co-ordinates. Better: Change the co-ordinate system to one that is more suitable to a disk.

Polar co-ordinates: $dx\, dy = r\, dr\, d\phi$.

$$\begin{aligned}
\iint\limits_{\text{unit disk}} (x^2 + y^2) dx\, dy &= \iint\limits_{\text{unit disk}} r^3 dr\, d\phi \\
&= \int_0^{2\pi} \int_0^1 r^3 dr\, d\phi \\
&= \int_0^{2\pi} \frac{1}{4} [r^3]_0^1\, d\phi \\
&= \int_0^{2\pi} \frac{1}{4} 1\, d\phi \\
&= \frac{1}{4} [\phi]_0^{2\pi} \\
&= \frac{\pi}{2}
\end{aligned}$$

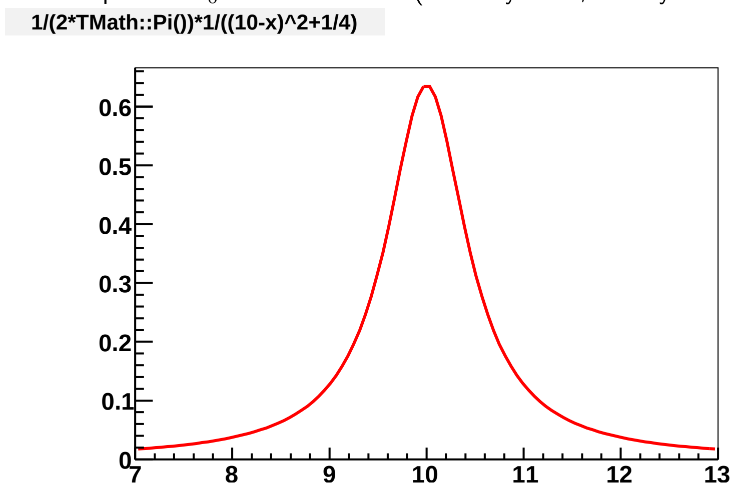
3.3 Additional Statistics Questions

3.3.1 Breit Wigner

The Breit Wigner function describes the probability density to find that a short-lived particle with mean mass M_0 and lifetime $\tau = \frac{\hbar}{\Gamma c^2}$ actually has the mass M.

$$\mathrm{B}_{\mathrm{M}_0}(M) = \frac{1}{2\pi}\frac{\Gamma}{(M-M_0)^2 + (\Gamma/2)^2}$$

An example for $M_0 = 10$ and $\Gamma = 1$ (arbitrary units, usually one would use MeV or GeV), is given below:



i) Show that the maximum of the Breit Wigner is at $M = M_0$

$$\begin{aligned}
\frac{dB}{dM} &= -\frac{1}{2\pi}\frac{\Gamma}{\left((M-M_0)^2 + (\Gamma/2)^2\right)^2}\cdot 2\,(M-M_0) \\
\left.\frac{dB}{dM}\right|_{M=M_{\max}} &= 0 \\
\Leftrightarrow (M-M_0) &= 0 \\
\Leftrightarrow M_{\max} = M_0 &
\end{aligned}$$

We won't bother with the 2nd derivative here, because it is pretty obvious that this is a maximum.

ii) Show that the FWHM of the Breit Wigner is Γ

$$\begin{aligned}
B(M_{\max}) &= \frac{1}{2\pi}\frac{\Gamma}{\Gamma^2/4} \\
B(M_{\max} + \frac{1}{2}\Gamma) = B(M_{\max} - \frac{1}{2}\Gamma) &= \frac{1}{2\pi}\frac{\Gamma}{\Gamma^2/4 + \Gamma^2/4} = \frac{1}{2}B(M_{\max}) \quad QED
\end{aligned}$$

Note that the Breit Wigner function neither has a well-defined mean, nor a standard deviation.