Problems Sheet: Statistics

MPAGS

1) χ^2 **fit** Download and complete the following jupyter notebook: `https://tinyurl.com/Chi2Fit`.

2) **MC integration** Download and complete the following jupyter notebook: `https://tinyurl.com/MCIntegration`.

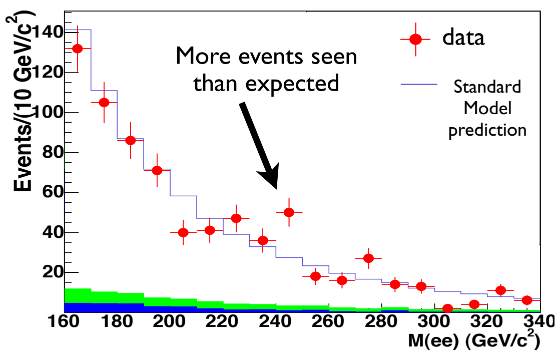
3) **Micromorts** Read about micromorts, here `https://plus.maths.org/content/os/issue55/features/risk/index` and then answer the questions I pondered on a ~ 1500 mile motorbike trip to Germany and back (during the few weeks in summer 2020 when this was possible), here: `https://youtu.be/qlu1coWbic8`. If you get stuck, here `https://tinyurl.com/MicromortsMCTrip` is a numerical solution which might point you in the right direction. But I'm looking for an analytic solution.

4) **Outbreak** Assume 0.2% of the population in the UK carry SARS-CoV-2 (Corona virus) (which is approximately the correct number). A test is developed that has a 99% chance of detecting the virus in a person who has it and a 1.5% chance of falsely indicating it in a person that does not have it. Calculate the probability that a person that tests positive for the virus, does indeed carry the virus. For an attempt of a visualisation of this problem, have a look at this jupyter notebook: `https://tinyurl.com/StatsOutbreak` .

5) **Which distribution?** Which distribution best describes the following:

   a) The number of flashes of lightening within one hour of a thunderstorm.
   b) The number of Higgs events at the LHC in a year of running.
   c) The number of students per hundred carrying the H1F1 virus.
   d) Weight of individual A4 pieces of paper in a notebook
   e) The number of sand grains in 1kg of sand.

6) $\mathbf{Z}' \to \mu\mu$**?** The plot below shows recent data from the CDF experiment at Fermilab:



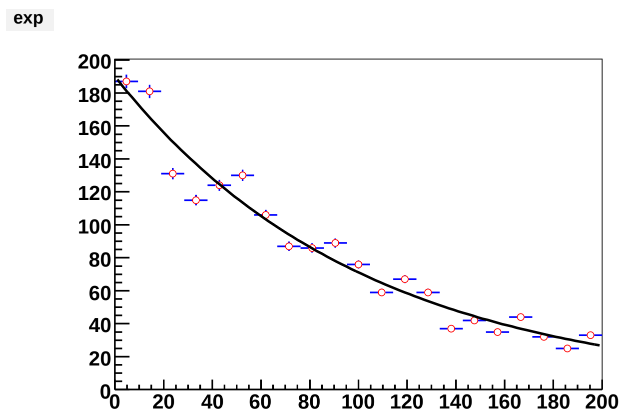
It displays the number of events in bins of the reconstructed mass for $Z' \to e^+e^-$ candidate. The Z' is a hypotetical partical, essentially a heavy Z, that pops up in a variety of beyond-the-Standard Model theories. The histogram (thin line) shows the expected number of events according to th the Standard Model. The dots with error bars represent the data. There appear to be more events than expected around $240\,\mathrm{GeV/c^2}$: 48 events where we expect 28 (note: I read these numbers off the graph, they are not exact).

(a) Estimate the significance of this observation, by calculating the probability of observing such an increased number of events as a result of a statistical fluctuation, taking into account that the physicists are looking simultaneously in 84 bins in the mass range from $160\,\mathrm{GeV/c^2}$ to $1000\,\mathrm{GeV/c^2}$ (for clarity, only some of that range is shown in the plot above).

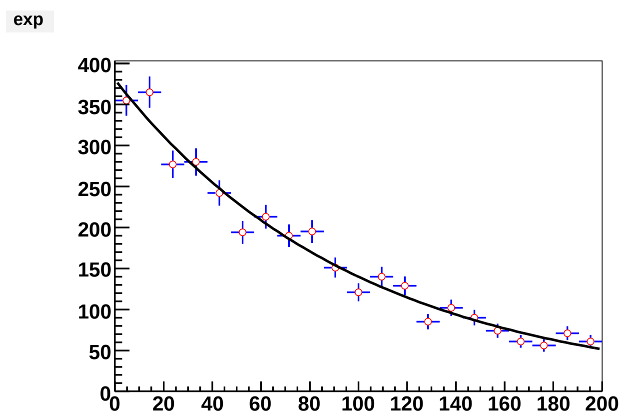
(b) Your answer above should be a fairly small number - less than a percent. So, this is quite an interesting result. Why do you think this has still not been published as a discovery of a new particle (that would have been all over the media)?

7) **Errorbars** Which of these 4 plots look like what you would expect if the theoretical prediction (black line) describes the data (red circles with blue error bars) and the error bars are Gaussian? Which ones don't? Why?

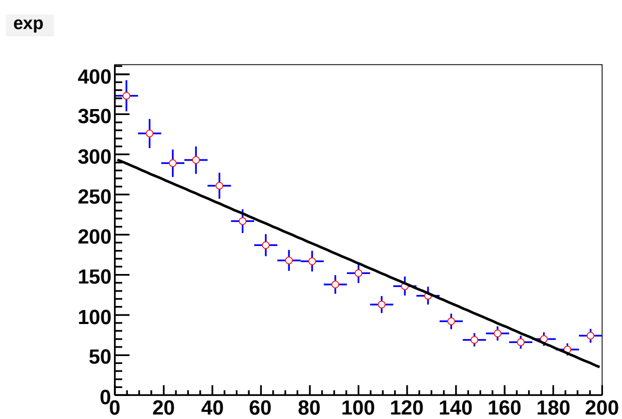
a)


exp
0
20
40
60
80
100
120
140
160
180
200
0
20
40
60
80
100
120
140
160
180
200

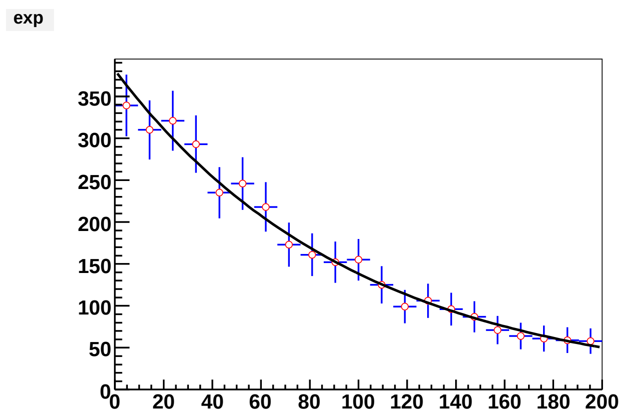
b)


exp
0
50
100
150
200
250
300
350
400
0
20
40
60
80
100
120
140
160
180
200

c)


exp
0
50
100
150
200
250
300
350
400
0
20
40
60
80
100
120
140
160
180
200

d)



8) Probability questions with a twist

   (i) When tossing a coin which has an equal chance of giving you heads or tails, which of the following (ordered) sequences is the most likely?

   (A) HHHHH

   (B) TTHHHHH

   (C) TTHHHHHTT

   (ii) In a terrible road accident on a usually save and accident-free stretch of road in South Yorkshire, 5 people get severely injured. Rank the following statements by the probability that they are true.

   (A) The accident was caused by an experienced driver.

   (B) The accident was caused by an experienced driver who was drunk.

9) The probability density for an unstable particle that exists at time $t_0 = 0$ to decay at a later time t is given by

$$P(t) = \left\{ \begin{array}{cc} Ke^{-t/\tau} & \text{if } t > 0 \\ 0 & \text{if else} \end{array} \right\}$$

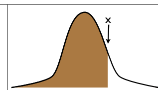
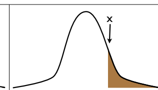
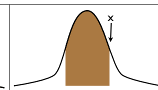
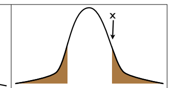
   (a) Show that $K = \frac{1}{\tau}$

   (b) What is the expectation value and the standard deviation of $P(t)$?

   (c) There are 3 particles at $t = 0$. What is the probability that at least one of these particles still exists (i.e. has not decayed) at $t = 3\tau$?

   (d) Given the PDF for the variable t given above, find the PDF for the variable $x(t) = 1 - e^{-t/\tau}$.

Integals over Gaussians

				
$s = \dfrac{x-\mu}{\sigma}$	$\int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$	$1 - \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$	$\int_{-s}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$	$1 - \int_{-s}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$
0	0.500	5.0E-01	0.000	1.0E+00
0.1	0.540	4.6E-01	0.080	9.2E-01
0.2	0.579	4.2E-01	0.159	8.4E-01
0.3	0.618	3.8E-01	0.236	7.6E-01
0.4	0.655	3.4E-01	0.311	6.9E-01
0.5	0.691	3.1E-01	0.383	6.2E-01
0.6	0.726	2.7E-01	0.451	5.5E-01
0.7	0.758	2.4E-01	0.516	4.8E-01
0.8	0.788	2.1E-01	0.576	4.2E-01
0.9	0.816	1.8E-01	0.632	3.7E-01
1	0.841	1.6E-01	0.683	3.2E-01
1.1	0.864	1.4E-01	0.729	2.7E-01
1.2	0.885	1.2E-01	0.770	2.3E-01
1.3	0.903	9.7E-02	0.806	1.9E-01
1.4	0.919	8.1E-02	0.838	1.6E-01
1.5	0.933	6.7E-02	0.866	1.3E-01
1.6	0.945	5.5E-02	0.890	1.1E-01
1.7	0.955	4.5E-02	0.911	8.9E-02
1.8	0.964	3.6E-02	0.928	7.2E-02
1.9	0.971	2.9E-02	0.943	5.7E-02
2	0.977	2.3E-02	0.954	4.6E-02
2.1	0.982	1.8E-02	0.964	3.6E-02
2.2	0.986	1.4E-02	0.972	2.8E-02
2.3	0.9893	1.1E-02	0.9786	2.1E-02
2.4	0.9918	8.2E-03	0.9836	1.6E-02
2.5	0.9938	6.2E-03	0.9876	1.2E-02
2.6	0.9953	4.7E-03	0.9907	9.3E-03
2.7	0.9965	3.5E-03	0.9931	6.9E-03
2.8	0.9974	2.6E-03	0.9949	5.1E-03
2.9	0.9981	1.9E-03	0.9963	3.7E-03
3	0.99865	1.3E-03	0.99730	2.7E-03
3.1	0.99903	9.7E-04	0.99806	1.9E-03
3.2	0.99931	6.9E-04	0.99863	1.4E-03
3.3	0.99952	4.8E-04	0.99903	9.7E-04
3.4	0.99966	3.4E-04	0.99933	6.7E-04
3.5	0.99977	2.3E-04	0.99953	4.7E-04
3.6	0.999841	1.6E-04	0.999682	3.2E-04
3.7	0.999892	1.1E-04	0.999784	2.2E-04
3.8	0.999928	7.2E-05	0.999855	1.4E-04
3.9	0.999952	4.8E-05	0.999904	9.6E-05
4	0.999968	3.2E-05	0.999937	6.3E-05