Lab 1 Physics 434

Ondrea Robinson

October 2019

1 A Little Statistics

1.1 Converting a Probability into a Sigma

1.1.1

The normal or Gaussian distribution is the standard distribution for most probabilistic data. It is a bell-shaped curve that centers around an average value with the highest incidence, with values far from the average decreasing in incidence.

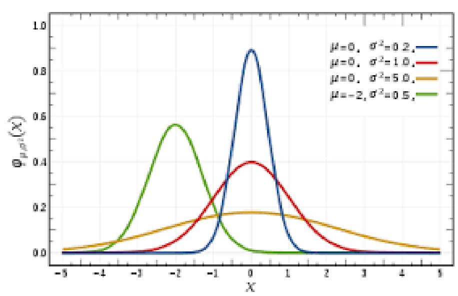

Figure 1: Graphs of three different normal PDFs

1.1.2

When we integrate the standard normal distribution, we have to integrate between two chosen points to find the probability that a measurement will be chosen between these two points. This is called the Cumulative Distribution Function, or CDF The CDF function is defined so that

$$cdf(b) = \int_{-\infty}^{b} pdf(x)dx$$

. Where pdf is the Probability Density Function defining the distribution we're looking at. In python, we can use the stats.norm.cdf() function:

```
import scipy
from scipy import stats
z = 1
table = stats.norm.cdf(z)
print(table)
```

From this code for values $z = [1, 3.5, 4]$ we get $table = []$. This matches the Z-table found on Wikipedia.

1.1.3

In this example we calculate the inverse function of the CDF called the Point Percentage Function. Instead of entering a sigma value we enter a probability and are returned the associated sigma. To show this, can enter the value given by the cdf function back into the ppf function and receive the original values back. To do this in python we use the code:

```
prob = stats.norm.ppf(table)
```

So that for $table = []$ it returns $prob = []$.

1.1.4

The ppf function returns a sigma value that is centered around a mean of zero. So if the sigma value returned is negative it just means the sigma associated with that probability is on the left of the mean.

1.2 The Lognormal Distribution

1.2.1

In the lognormal distribution, the data variable x is distributed such that if you take $log(x) = y$, y is normally distributed. This distribution is fairly common and used to model the time it takes to solve a Rubik's cube, the size of living tissues and city sizes. The lognormal PDF is defined so that

$$pdf(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp(-\frac{(\ln(x-\mu))^2}{2\sigma^2})$$

.

In python, we can graph the lognormal distribution with different parameters. The parameter s in this case defines sigma for the underlying normal distribution. The python code for the plots is:

```
import scipy
from scipy import stats
z = 1
table = stats.norm.cdf(z)
print(table)
```