**Analysis of Uncertainties** **Fall 2014**

# Lecture 2: 24 October 2014

*Lecturer: Prof. Kate Scholberg* *Notes by: Douglas Davis*

## 2.1 Probability

We define probability with

$$P = \frac{\text{number of occasions some particular result occurs}}{\text{total number of occurrences}} \tag{2.1}$$

The probability can only be between 0 and 1 (normalized). We write a discrete normalized distribution as:

$$\sum_i P_i = 1 \tag{2.2}$$

An example would be a the probability of rolling a dice and getting a certain number. If the dice is fair, we have 6 total outcomes with each having a probability of 1/6

$$\sum_{i=1}^{6} P_i = 6 \times \frac{1}{6} = 1$$

An example of continuous probabilities would be the isotropic emission of say a photon (let's take this to be 2-dimensional for simplicity). The probability that the photon comes off at a set angle $\theta$ is just

$$P(\theta) = \frac{1}{2\pi},$$

it's just constant for any angle. The probability that the photon is emitted in some range of angles, $\Delta\theta$, would be $\Delta\theta/2\pi$. Integrating over all of $2\pi$ angles, we show:

$$\int_0^{2\pi} P(\theta)d\theta = \frac{1}{2\pi}\int_0^{2\pi} d\theta = 2\pi/2\pi = 1.$$

The familiar normalization condition for wave functions in quantum mechanics is just a property of probability. The probability that the state represented by the wave function $\psi$ exists from $-\infty$ to $\infty$ is 1.

Let's develop some definitions and notation. If $A$ and $B$ are possible outcomes, then:

$$P(A+B) \equiv \text{probability of } A \text{ or } B, \tag{2.3}$$

When $A$ and $B$ are independent (exclusive) the $P(A+B)$ is just equal to the sum of $P(A)$ and $P(B)$, but in general:

$$P(A+B) \leq P(A) + P(B).$$

Another definition:

$$P(AB) \equiv \text{probability of } A \text{ and } B = P(A|B)P(B). \tag{2.4}$$

The new form we've introduced is the *conditional probability*, $P(A|B)$, which is the probability of $A$ given $B$. Let's jump into a couple examples using the conditional probability. Let's say that $A$ represents that it's a Friday in October, and $B$ represents we have a 771 class day during this day in October. There are 4 class days for 771 in October, and there are 5 Fridays in the current October (2 are on Friday, 2 are not). First let's determine if the $A$ and $B$ are independent (if they are, $P(A)P(B) = P(AB)$).

$$P(A)P(B) = \frac{5}{31} \times \frac{4}{31} \approx .02$$
$$P(AB) = P(B|A)P(A) = \frac{2}{4} \times \frac{4}{31} = \frac{2}{31} \approx .065.$$

Therefore, $A$ and $B$ are definitely not independent (as expected). We can also calculate $P(AB)$ using a different ordering of equation 2.4:

$$P(AB) = P(B|A)P(A) = \frac{2}{5} \times \frac{5}{31} = \frac{2}{31},$$

as expected.

We've now reached a point where we can introduce Bayes' Theorem:

$$P(A|B)P(B) = P(B|A)P(A) \tag{2.5}$$
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.6}$$

Taking an example that works perfectly for current events, let $A$ be the probability that you have ebola and let $B$ be the probability that you test positive. Let's say a test has a 99% detection rate and a 1% false detection rate.

$$P(B|A) = \text{testing positive if you have ebola (99\%)}$$
$$P(A|B) = \text{probability of having ebola if you test positive (unknown)}$$

We now need a value for $P(A)$, let's call the probability of having ebola 0.001. $P(B)$ is our false positive rate of 1%. Now we can calculate:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.99 \times 0.001}{0.01} \approx 10\%.$$

With this we something a little counter intuitive – if you test positive for ebola, there's really only a 10% chance you have it under these circumstances. Let's look at a better physics example, proton decay. Let $A$ be the probability of proton decay, $B$ is the probability of any event passing selection cuts.

$$P(B|A) = \text{effeciency (probability the event passes cuts given proton decay)}$$
$$P(A|B) = \text{probability of proton decay given an event passing selection cuts}$$

Using Bayes':

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{2.7}$$

where $P(B)$ is the probability of a background event – therefore the experiment needs high efficiency and low background to make a serious measurement (increase the value of $P(A|B)$).

Let's now consider a parameter estimation scenario. Take $\frac{1}{2}(1 + \alpha \cos\theta)$. Let $A$ be the experimental result probability and $B$ be the model parameterized by $\alpha$. $P(B|A) = P(A|B)P(B)/P(A)$ is the probability of $\alpha$ given the data. In this case $P(B)$ is very important, (this term is called the prior). If the experimenter does not know $P(B)$ well (if one does not have a reasonable distribution for it) then one cannot make much sense of $P(B|A)$ with much confidence. This is where Bayes' theorem has a weakness.

## 2.2 Probability Distributions

### 2.2.1 Binomial

Let's imagine we have three dice – how often will we get 3 ones? This is pretty simple:

$$(1/6)^3 \approx 0.0046,$$

about half a percent. How about 2 ones? Now we have 3 different scenarios where this will happen, The first a second dice will be 1, the first and third dice will be 1, and the second and third dice will be 1. Therefore wehave:

$$3 \times (1/6)^2(5/6) \approx 6.9\%$$

In this case – a "success" is getting a 1. Now we define a probability distribution using the following parameters:

$$P(x; n, p) \tag{2.8}$$

where $x$ is the number of successes, $n$ is the number of trials, and $p$ is the probability of success in one trial. In the above case, our distribution would be $P(x; 3, 1/6)$, and it would be a **Binomial Distribution**. The probability distribution function for a binomial distribution is:

$$P(x; n, p) = \binom{n}{x} p^x q^{n-x}, \tag{2.9}$$

where

$$\binom{n}{x} \equiv n \text{ “choose” } x = \frac{n!}{x!\,(n-x)!} \text{ and } q \equiv 1 - p$$

Now we define the mean and the variance of the binomial distribution:

$$\langle x \rangle = \mu = \sum_{x=0}^{n} x P(x; n, p) = np \tag{2.10}$$

$$\sigma^2 = np(1-p) \tag{2.11}$$

### 2.2.2 Poisson

The limit of the binomial distribution for low rate processes ($\mu \ll n, p \ll 1$) yields the **Poisson distribution**:

$$P(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}. \tag{2.12}$$

Poisson is unique in that the expected value and variance are the same:

$$\langle x \rangle = \mu \tag{2.13}$$

$$\sigma^2 = \mu \tag{2.14}$$

Poisson is continuous in $\mu$ and discrete in $x$.

### 2.2.3 Gaussian

Now we take $np \gg 1$ and arrive at the **Gaussian distribution**:

$$P(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]. \tag{2.15}$$

The Gaussian is defined for all $x$ and is continuous. Some properties of the distribution:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} dx\, P(x;\mu,\sigma) \\ \langle x \rangle &= \int_{-\infty}^{\infty} dx\, x\, P(x;\mu,\sigma) \\ \langle (x-\mu)^2 \rangle &= \int_{-\infty}^{\infty} dx\, (x-\mu)^2\, P(x;\mu,\sigma) \\ P(\mu \pm \sigma;\mu,\sigma) &= \frac{1}{e} P(\mu;\mu,\sigma) \end{aligned}$$

The "full width at half max", $\Gamma$, is $2.354\sigma$. And finally:

$$\begin{aligned} 1\sigma \text{ out from mean } &\approx 68\% \text{ of area} \\ 2\sigma &\approx 95\% \text{ of area} \\ 3\sigma &\approx 99\% \text{ of area} \end{aligned}$$