Statistics, Probability Distributions & Error Propagation

James R. Graham
9/2/08
Sample & Parent Populations

• Make measurements
  − $x_1$
  − $x_2$
  − In general do not expect $x_1 = x_2$
  − But as you take more and more measurements a pattern emerges in this sample

• With an infinite sample $x_i$, $i \in \{1\ldots\infty\}$
  − Expect a pattern emerges with a characteristic value
  − Exactly specify the distribution of $x_i$
  − This hypothetical pool of all possible measurements is the parent population
  − Any finite sequence is the sample population
Histograms & Distributions

- Histogram represents the **frequency** of discrete measurements
  - True parent population (dotted)
  - Inferred parent distribution (solid)
Notation

• Parent distribution: Greek, e.g., $\mu$
• Sample distribution: Latin, $\bar{x}$
  – To determine properties of the parent distribution we assume that the properties of the sample distribution tend to those of the parent as $N$ tends to infinity
Summation

• If we make \( N \) measurements, \( x_1, x_2, x_3, \) etc. the sum of these measurements is

\[
\sum_{i=1}^{N} x_i = x_1 + x_2 + x_3 + \ldots + x_N
\]

• Typically, we use the shorthand

\[
\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i
\]
Mean

• The mean of an experimental distribution is

\[ \bar{x} = \frac{1}{N} \sum x_i \]

• The mean of the parent population is defined as

\[ \mu = \lim_{N \to \infty} \left( \frac{1}{N} \sum x_i \right) \]
Median

- The median of the parent population $\mu_{1/2}$ is the value for which half of $x_i < \mu_{1/2}$

$$P(x_i < \mu_{1/2}) = P(x_i \geq \mu_{1/2}) = 1/2$$

- The median cuts the area under the probability distribution in half
Mode

• The mode is the most probable value drawn from the parent distribution
  – The mode is the most likely value to occur in an experiment
  – For a symmetrical distribution the mean, median and mode are all the same
Deviations

• The deviation, \( d_i \), of a measurement, \( x_i \), from the mean, \( \mu \), is defined as

\[
d_i = x_i - \mu
\]

• If \( \mu \) is the true mean value the deviation is the error in \( x_i \)
Mean Deviation

• The mean deviation vanishes
  – Evident from the definition

\[ \lim_{N \to \infty} d = \lim_{N \to \infty} \left[ \frac{1}{N} \sum (x_i - \mu) \right] = \lim_{N \to \infty} \left[ \frac{1}{N} \sum x_i \right] - \mu \]
Mean Square Deviation

• The mean square deviation is easy to use analytically and justified theoretically

\[ \sigma^2 = \lim_{N \to \infty} \left[ \frac{1}{N} \sum (x_i - \mu)^2 \right] = \lim_{N \to \infty} \left[ \frac{1}{N} \sum x_i^2 \right] - \mu^2 \]

• \( \sigma^2 \) is also known as the **variance**
  – Derive this expression
  – Computation of \( \sigma^2 \) assumes we know \( \mu \)
Population Mean Square Deviation

• The standard deviation, $s$, of a sample population is given by

$$s^2 = \frac{1}{N-1} \sum (x_i - \bar{x})^2$$

• The factor $(N-1)$ rather than $N$ accounts for the fact that the mean was derived from the data
Significance

• The mean of the sample is the best estimate of the mean of the parent distribution
  – The standard deviation, $s$, is characteristic of the uncertainties associated with attempts to measure $\mu$
  – But what is the uncertainty in $\mu$?
• To answer these questions we need probability distributions…
\( \mu \) and \( \sigma \) of Distributions

- Define \( \mu \) and \( \sigma \) in terms of the parent probability distribution \( P(x) \)
  - Definition of \( P(x) \)
    - Limit as \( N \to \infty \)
    - The number of observations \( dN \) that yield values between \( x \) and \( x + dx \) is
      \[
      \frac{dN}{N} = P(x) \, dx
      \]
Expectation Values

• The mean, $\mu$, is the expectation value
  
  $<x>$

• The variance, $\sigma^2$, is the expectation value
  
  $<(x-\mu)^2>$
**Expectation Values**

- For a discrete distribution, $N$, observations and $n$ distinct outcomes

\[
\mu = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} x_i
\]

\[
= \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{n} [x_j NP(x_j)]
\]

\[
= \lim_{N \to \infty} \sum_{j=1}^{n} [x_j P(x_j)]
\]
Expectation Values

• For a discrete distribution, \( N \), observations and \( n \) distinct outcomes

\[
\sigma^2 = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2
= \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{n} (x_j - \mu)^2 NP(x_j)
= \lim_{N \to \infty} \sum_{j=1}^{n} ((x_j - \mu)^2 P(x_j))
\]
Expectation values

- The expectation value of any continuous function of $x$

\[
\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) P(x) \, dx
\]
\[
\mu = \int_{-\infty}^{\infty} x P(x) \, dx
\]
\[
\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 P(x) \, dx
\]

where \( \int_{-\infty}^{\infty} P(x) \, dx = 1 \)
Binomial Distribution

• Suppose we have two possible outcomes with probability $p$ and $q = 1 - p$
  – e.g., a coin toss, $p = 1/2$, $q = 1/2$
• If we flip $n$ coins what is the probability of getting $x$ heads?
  – Answer is given by the Binomial Distribution

\[
P(x; n, p) = C(n, x) p^x q^{n-x}
\]

– $C(n, x)$ is the number of combinations of $n$ items taken $x$ at a time $= n!/[x!(n-x)!]$
**Binomial Distribution**

- The expectation value

\[
\mu = \sum_{x=0}^{n} x P(x; n, p) = \sum_{x=0}^{n} x C(n, x) p^x q^{n-x} = \sum_{x=0}^{n} \left[ x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np
\]
Poisson Distribution

• The Poisson distribution is the limit of the Binomial distribution when $\mu \ll n$ because $p$ is small
  – The binomial distribution describes the probability $P(x; n, p)$ of observing $x$ events per unit time out of $n$ possible events
  – Usually we don’t know $n$ or $p$ but we do know $\mu$


**Poisson Distribution**

- Suppose $p \ll 1$ then $x \ll n$

\[ P(x;n,p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \]

\[ \frac{n!}{x!(n-x)!} = n(n-1)(n-2)...(n-x-2)(n-x-1) \approx n^x \text{ when } n \gg x \]

\[ \frac{n!}{(n-x)!} p^x \approx (np)^x = \mu^x \]

\[ (1-p)^{n-x} = (1-p)^{-x}(1-p)^n \approx 1 \times (1-p)^n \text{ since } p \ll 1 \]

\[ \lim_{p \to 0} (1-p)^n = \lim_{p \to 0} [(1-p)^{1/p}]^\mu = (e^{-1})^\mu = e^{-\mu} \]

\[ P(x,\mu) = \frac{\mu^x}{x!} e^{-\mu} \]
Poisson Distribution

• The expectation value of $x$ is

$$\langle x \rangle = \sum_{x=0}^{\infty} x P(x, \mu) = \sum_{x=0}^{\infty} x \frac{\mu^x}{x!} e^{-\mu} = \mu$$

• Expectation value of $(x-\mu)^2$

$$\sigma^2 = \langle (x - \mu)^2 \rangle = \sum_{x=0}^{\infty} (x - \mu)^2 \frac{\mu^x}{x!} e^{-\mu} = \mu$$
Gaussian or Normal Distribution

• The Gaussian distribution is an approximation to the binomial distribution for large \( n \) and large \( np \)

\[
P(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2}
\]
Gaussian or Normal Distribution

\[ P(x; \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} \]

\[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \, dx = 0.683 \]

\[ +/- 1\sigma: 68.3\% \]
\[ +/- 2\sigma: 95.5\% \]
\[ +/- 3\sigma: 99.7\% \]
Combining Two Observations

• Suppose I have two sets of measurements, \( a_i \) and \( b_i \)
  – A derived quantity \( c_i = a_i + b_i \)
  – What is the relation between the means and standard deviations of \( a_i \) and \( b_i \) and \( c_i \)
  – Suppose we have the same number of observations \( N \)
Combining Two Observations

\[ N = N_a = N_b \]

\[ \bar{a} = \frac{1}{N} \sum a_i \quad \bar{b} = \frac{1}{N} \sum b_i \]

\[ \bar{c} = \frac{1}{N} \sum c_i \quad s_c^2 = \frac{1}{N-1} \sum (c_i - \bar{c})^2 \]

\[ c_i = a_i + b_i \]

\[ \bar{c} = \frac{1}{N} \sum (a_i + b_i) = \frac{1}{N} \sum a_i + \frac{1}{N} \sum b_i \]

\[ = \bar{a} + \bar{b} \]
Combining Two Observations

\[ s_c^2 = \frac{1}{N-1} \sum (c_i - \bar{c})^2, \quad \bar{c} = \bar{a} + \bar{b} \]

\[ s_c^2 = \frac{1}{N-1} \sum \left[ a_i + b_i - (\bar{a} + \bar{b}) \right]^2 \]

\[ = \frac{1}{N-1} \sum \left[ (a_i + b_i)^2 - 2(a_i + b_i)(\bar{a} + \bar{b}) + (\bar{a} + \bar{b})^2 \right] \]

\[ = \frac{1}{N-1} \sum \left[ a_i^2 + b_i^2 + 2a_i b_i - 2(a_i \bar{a} + a_i \bar{b} + b_i \bar{a} + b_i \bar{b}) + (\bar{a})^2 + 2\bar{a} \bar{b} + (\bar{b})^2 \right] \]

\[ = \frac{N}{N-1} \bar{a}^2 + \frac{N}{N-1} \bar{b}^2 + \frac{2}{N-1} \sum a_i b_i - \frac{N}{N-1} (\bar{a})^2 - \frac{2N}{N-1} \bar{a} \bar{b} - \frac{N}{N-1} (\bar{b})^2 \]
Combining Two Observations

\[ s_c^2 = \frac{1}{N-1} \sum (c_i - \bar{c})^2, \quad \bar{c} = \bar{a} + \bar{b} \]

\[ = \frac{N}{N-1} \overline{a^2} + \frac{N}{N-1} \overline{b^2} + \frac{2}{N-1} \sum a_i b_i - \frac{N}{N-1} (\overline{a})^2 - \frac{2N}{N-1} \overline{a} \overline{b} - \frac{N}{N-1} (\overline{b})^2 \]

\[ = \frac{N}{N-1} \left[ \overline{a^2} - (\overline{a})^2 \right] + \frac{N}{N-1} \left[ \overline{b^2} - (\overline{b})^2 \right] + \frac{2N}{N-1} \left( \overline{ab} - \overline{a} \overline{b} \right) \]

\[ s_c^2 = s_a^2 + s_b^2 + 2s_{ab}^2 \]

- The term \( s_{ab}^2 \) is the covariance
  - Murphy’s law factor
  - \( s_{ab} \) can be negative, zero or positive
Combining Two Uncorrelated Observations

• When $a$ and $b$ are uncorrelated the covariance is zero

\[ s_{ab}^2 = \frac{1}{N-1} \sum (a_i - \bar{a})(b_i - \bar{b}) = 0 \]

\[ s_c^2 = s_a^2 + s_b^2 \]

– The variance of $c$ is the sum of the variances of $a$ and $b$

• This demonstrates the fundamentals of error propagation
**Propagation of Errors**

- Suppose we want to determine $x$ which is a function of measured quantities, $u$, $v$, etc.

  $$x = f(u, v, ...)$$

- Assume that

  $$\bar{x} = f(\bar{u}, \bar{v}, ...)$$
**Propagation of Errors**

- The uncertainty in $x$ can be found by considering the spread of the values of $x$ resulting from individual measurements, $u_i$, $v_i$, etc.,

$$x_i = f(u_i, v_i, \ldots)$$

- In the limit of $N \to \infty$ the variance of $x$

$$\sigma_x^2 = \lim_{N \to \infty} \frac{1}{N} \sum_{i} (x_i - \bar{x})^2$$
Propagation of Errors

- Taylor expand the deviation

\[ x_i - \bar{x} = (u_i - \bar{u}) \left. \frac{\partial f}{\partial u} \right|_{\bar{u}} + (v_i - \bar{v}) \left. \frac{\partial f}{\partial v} \right|_{\bar{v}} + ... \]

\[ \sigma_x^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i \left[ (u_i - \bar{u}) \left. \frac{\partial f}{\partial u} \right|_{\bar{u}} + (v_i - \bar{v}) \left. \frac{\partial f}{\partial v} \right|_{\bar{v}} + ... \right]^2 \]

\[ = \lim_{N \to \infty} \frac{1}{N} \sum_i \left[ (u_i - \bar{u})^2 \left( \left. \frac{\partial f}{\partial u} \right|_{\bar{u}} \right)^2 + (v_i - \bar{v})^2 \left( \left. \frac{\partial f}{\partial v} \right|_{\bar{v}} \right)^2 + 2(u_i - \bar{u})(v_i - \bar{v}) \left. \frac{\partial f}{\partial u} \right|_{\bar{u}} \left. \frac{\partial f}{\partial v} \right|_{\bar{v}} + ... \right] \]
Propagation of Errors

\[ \sigma_x^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i \left[ (u_i - \bar{u})^2 \left( \frac{\partial f}{\partial u} \right)^2_{\bar{u}} + (v_i - \bar{v})^2 \left( \frac{\partial f}{\partial v} \right)^2_{\bar{v}} + 2(u_i - \bar{u})(v_i - \bar{v}) \frac{\partial f}{\partial u} \bigg|_{\bar{u}} \frac{\partial f}{\partial v} \bigg|_{\bar{v}} \right] + \ldots \]

\[ = \lim_{N \to \infty} \frac{1}{N} \sum_i (u_i - \bar{u})^2 \left( \frac{\partial f}{\partial u} \right)^2_{\bar{u}} + \ldots \]

\[ = \left[ \lim_{N \to \infty} \frac{2}{N} \sum_i (u_i - \bar{u})(v_i - \bar{v}) \frac{\partial f}{\partial u} \bigg|_{\bar{u}} \frac{\partial f}{\partial v} \bigg|_{\bar{v}} \right] + \ldots \]

\[ \sigma_x^2 = \sigma_u^2 \left( \frac{\partial f}{\partial u} \right)^2_{\bar{u}} + \sigma_v^2 \left( \frac{\partial f}{\partial v} \right)^2_{\bar{v}} + 2\sigma_{uv}^2 \frac{\partial f}{\partial u} \bigg|_{\bar{u}} \frac{\partial f}{\partial v} \bigg|_{\bar{v}} + \ldots \]
Examples of Error Propagation

• Suppose $a = b + c$
  – We know that

\[
\bar{a} = \bar{b} + \bar{c}
\]

\[
\sigma_a^2 = \sigma_b^2 + \sigma_c^2
\]

assuming that the covariance is 0

• What about $a = b/c$?
Examples of Error Propagation

• Suppose \( a = b/c \)?

\[
\bar{a} = \frac{\bar{b}}{\bar{c}}
\]

and

\[
\sigma_a^2 = \sigma_b^2 \left( \frac{\partial a}{\partial b} \right)_b^2 + \sigma_c^2 \left( \frac{\partial f}{\partial v} \right)_c^2 + 2 \sigma_{bc} \left. \frac{\partial a}{\partial b} \right|_b \left. \frac{\partial f}{\partial c} \right|_c + \ldots
\]

\[
\sigma_a^2 = \sigma_b^2 \frac{1}{c^2} + \sigma_c^2 \left( \frac{b}{c^2} \right)^2
\]

or

\[
\frac{\sigma_a^2}{a^2} = \frac{\sigma_b^2}{b^2} + \frac{\sigma_c^2}{c^2}
\]

assuming that the covariance is 0
Examples of Error Propagation

• What happens when \( m = -2.5 \log_{10}(F/F_0) \)?
  – What is the error in \( m \)?

\[
m = -2.5 \log_{10}(F/F_0)
\]

and

\[
\sigma_m^2 = \sigma_F^2 \left( \frac{\partial m}{\partial F} \right)_F^2
\]

\[
\sigma_m^2 = \sigma_F^2 \left( \frac{2.5}{F \log(10)} \right)^2
\]

\[
\sigma_m^2 = \left( 1.087 \right)^2 \left( \frac{\sigma_F}{F} \right)^2
\]
Error of the Mean

• Suppose we have \( N \) measurements, \( x_i \) with uncertainties characterized by \( s_i \)

\[
\bar{x} = \frac{1}{N} (x_1 + x_2 + x_3 + \ldots + x_N) = \frac{1}{N} \sum_i x_i
\]

\[
s_{\bar{x}}^2 = s_1^2 \left( \frac{\partial \bar{x}}{\partial x_1} \right)^2 + s_2^2 \left( \frac{\partial \bar{x}}{\partial x_2} \right)^2 + s_3^2 \left( \frac{\partial \bar{x}}{\partial x_3} \right)^2 + \ldots + s_N^2 \left( \frac{\partial \bar{x}}{\partial x_N} \right)^2
\]

\[
= \sum_i s_i^2 \left( \frac{\partial \bar{x}}{\partial x_i} \right)^2
\]
Suppose the errors on all points are equal so that \( s_i = s \)

\[
\begin{align*}
    s_{\bar{x}}^2 &= \sum_i s_i^2 \left( \frac{\partial \bar{x}}{\partial x_i} \right)^2 \\
    \frac{\partial \bar{x}}{\partial x_i} &= \frac{\partial}{\partial x_i} \left( \frac{1}{N} \sum_j x_j \right) = \frac{1}{N} \\
    s_{\bar{x}}^2 &= \sum_i s_i^2 \left( \frac{1}{N} \right)^2 \\
    &= \frac{s^2}{N}
\end{align*}
\]