

Earth Science Applications of Space Based Geodesy

DES-7355

Tu-Th

9:40-11:05

Seminar Room in 3892 Central Ave. (Long building)

Bob Smalley

Office: 3892 Central Ave, Room 103

678-4929

Office Hours – Wed 14:00-16:00 or if I'm in my office.

http://www.ceri.memphis.edu/people/smalley/ESCI7355/ESCI_7355_Applications_of_Space_Based_Geodesy.html

Class 7

More inversion pitfalls

Bill and Ted's misadventure.

Bill and Ted are geo-chemists who wish to measure the number of grams of each of three different minerals A, B, C held in a single rock sample.

Let

a be the number of grams of A ,

b be the number of grams of B ,

c be the number of grams of C

d be the number of grams in the sample.

By performing complicated experiments Bill and Ted are able to measure four relationships between a, b, c, d which they record in the matrix below:

$$\begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \\ 23.2134 & -86.3925 & 44.693 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 34.7177 \\ 70.9241 \\ 82.9271 \\ -26.222 \end{pmatrix}$$

$$Ax = b$$

Now we have more equations than we need
What to do?

One thing to do is throw out one of the equations

(in reality only a Mathematician is naïve enough to think that three equations is sufficient to solve for three unknowns – but lets try it anyway).

So throw out one – leaving



(different A and b from before)

Remembering some of their linear algebra they know that the matrix is not invertible if the determinant is zero, so they check that

$$\begin{vmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{vmatrix} \approx -2$$

OK so far
(or “fat, dumb and happy”)

So now we can compute



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

So now we're done.

Or are we?

Next they realize that the measurements are really only
good to 0.1

So they round to 0.1 and do it again



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Now they notice a small problem –

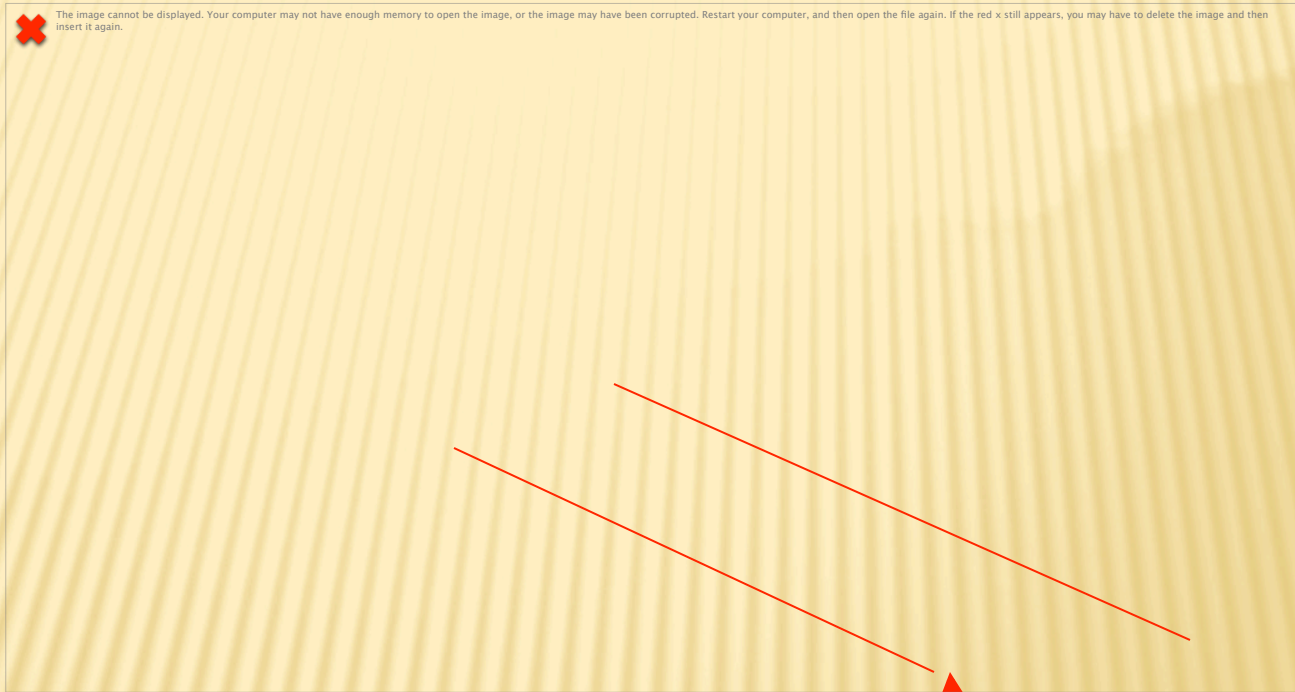
They get a very different answer

(and they don't notice they have a bigger physical problem in that they have negative weights/amounts!)

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.8 \\ 0.7 \end{pmatrix}$$

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1.68294 \\ 8.92282 \\ -3.50254 \end{pmatrix}$$

So what's the problem?
First find the SVD of A .



Since there are three non-zero values on the diagonal A is invertible

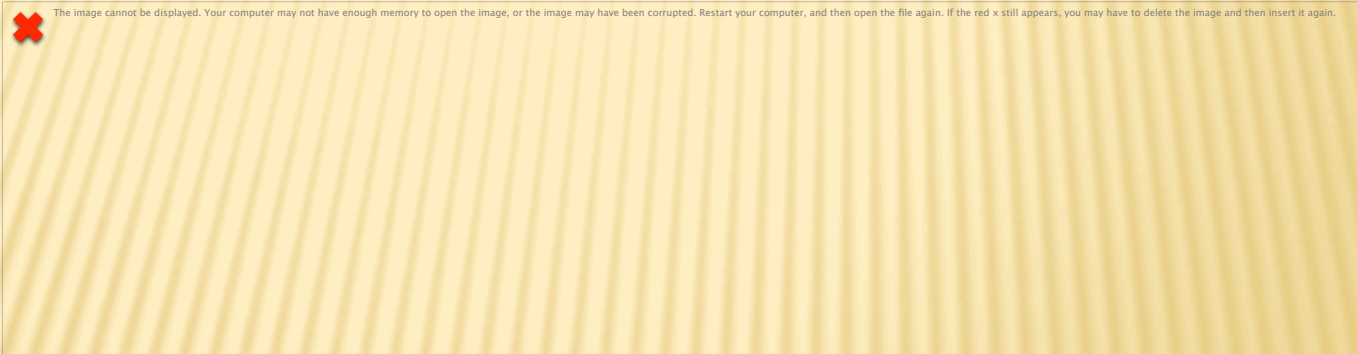
BUT, one of the singular values is much, much less than the others

$$A = \begin{pmatrix} 93.477 & 10.202 & -28.832 \\ 1.93 & 32.816 & 62.414 \\ 26.821 & 36.816 & 57.234 \end{pmatrix} =$$

$$A = (\vec{h}_1 \ \vec{h}_2 \ \vec{h}_3) \begin{pmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 0.0002 \end{pmatrix} \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{pmatrix}$$

So the matrix is “almost” rank 2
(which would be non-invertible)

We can also calculate the SVD of A^{-1}



So now we can see what happened
(why the two answers were so different)

Let y be the first version of b

Let y' be the second version of b (to 0.1)

$$\begin{aligned} |A^{-1}y - A^{-1}y'| &= |A^{-1}(y - y')| = \\ &= \left| \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{pmatrix} \begin{pmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 5000 \end{pmatrix} \begin{pmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vec{h}_3 \end{pmatrix} (y - y') \right| \end{aligned}$$

So A^{-1} stretches vectors parallel to h_3 and a_3 by a factor of 5000.

Returning to GPS



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

We have 4 unknowns (x^R, y^R, z^R and τ^R)
And 4 (nonlinear) equations
(later we will allow more satellites)
So we can solve for the unknowns

Again, we cannot solve this directly

Will solve iteratively by

- 1) Assuming a location
- 2) Linearizing the range equations
- 3) Use least squares to compute new (better) location
- 4) Go back to 1 using location from 3

We do this till some convergence criteria is met (if we're lucky)

linearize

So - for one satellite we have



Linearize (first two terms of Taylor Series)

$$P(x, y, z, \tau) \approx P(x_0, y_0, z_0, \tau_0) + (x - x_0) \frac{\partial P}{\partial x} \bigg|_{(x_0, y_0, z_0, \tau_0)} \\ + (y - y_0) \frac{\partial P}{\partial y} \bigg|_{(x_0, y_0, z_0, \tau_0)} + (z - z_0) \frac{\partial P}{\partial z} \bigg|_{(x_0, y_0, z_0, \tau_0)} + (\tau - \tau_0) \frac{\partial P}{\partial \tau} \bigg|_{(x_0, y_0, z_0, \tau_0)}$$

$$P(x, y, z, \tau) \approx P_{\text{computed}} + \frac{\partial P}{\partial x} \Delta x + \frac{\partial P}{\partial y} \Delta y + \frac{\partial P}{\partial z} \Delta z + \frac{\partial P}{\partial \tau} \Delta \tau$$

Residual

Difference between observed and calculated
(linearized)



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

So we have the following for one satellite

$$\Delta P = \frac{\partial P}{\partial x} \Delta x + \frac{\partial P}{\partial y} \Delta y + \frac{\partial P}{\partial z} \Delta z + \frac{\partial P}{\partial \tau} \Delta \tau + \vec{v}$$

Which we can recast in matrix form

$$\Delta P = \begin{pmatrix} \frac{\partial P}{\partial x} & \frac{\partial P}{\partial y} & \frac{\partial P}{\partial z} & \frac{\partial P}{\partial \tau} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{pmatrix} + \vec{v}$$

For m satellites (where $m \geq 4$)



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Which is usually written as



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Calculate the derivatives



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

So we get

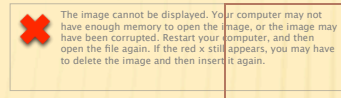
$$A = \begin{pmatrix} \frac{x_0 - x^1}{\rho^1} & \frac{y_0 - y^1}{\rho^1} & \frac{z_0 - z^1}{\rho^1} & c \\ \frac{x_0 - x^2}{\rho_2} & \frac{y_0 - y^2}{\rho_2} & \frac{z_0 - z^2}{\rho_2} & c \\ \frac{x_0 - x^3}{\rho_3} & \frac{y_0 - y^3}{\rho_3} & \frac{z_0 - z^3}{\rho_3} & c \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_0 - x^m}{\rho_m} & \frac{y_0 - y^m}{\rho_m} & \frac{z_0 - z^m}{\rho_m} & c \end{pmatrix}$$

Is function of direction to satellite

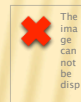
Note last column is a constant

Consider some candidate solution x'

Then we can write

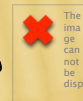


b are the observations



are the residuals

We would like to find the x' that minimizes



So the question now is how to find this x'

One way, and the way we will do it,

Least Squares

Since we have already done this – we'll go fast

Use solution to linearized form of observation equations to write estimated residuals

$$\vec{v} = \vec{b} - A\hat{x}'$$

Vary value of x to minimize

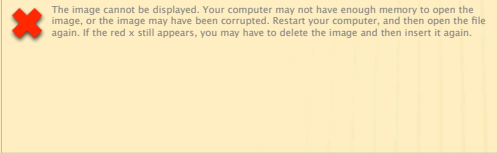
$$J(\vec{x}) = \sum_{i=1}^m \vec{v}_i^2 = \vec{v}^T \vec{v} = (\vec{b} - A\vec{x})^T (\vec{b} - A\vec{x})$$



The image cannot be displayed. Your computer may not have enough memory to open the image, or the image may have been corrupted. Restart your computer, and then open the file again. If the red x still appears, you may have to delete the image and then insert it again.

Normal equations

Solution to normal equations



Assumes

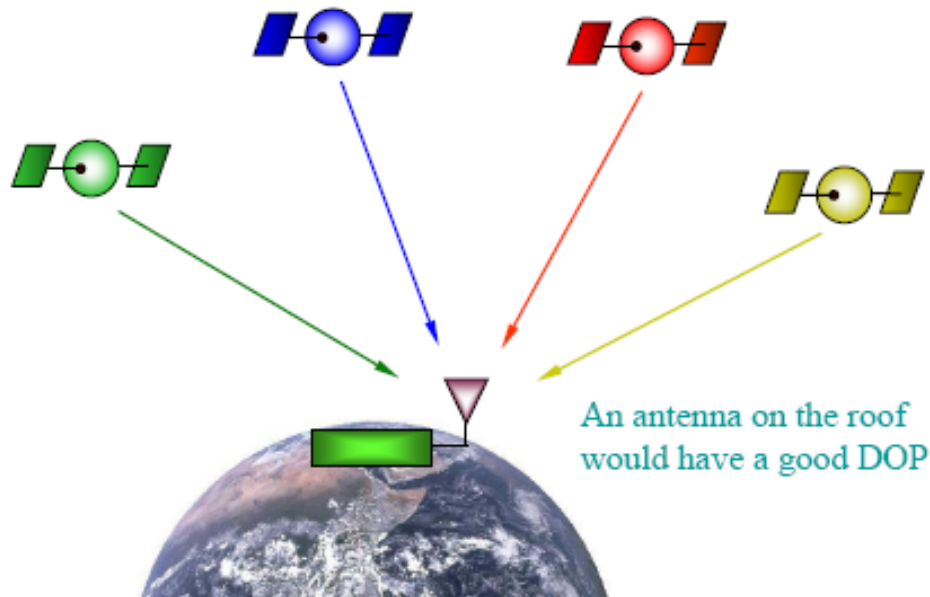
Inverse exists

(m greater than or equal to 4, necessary but not sufficient condition)

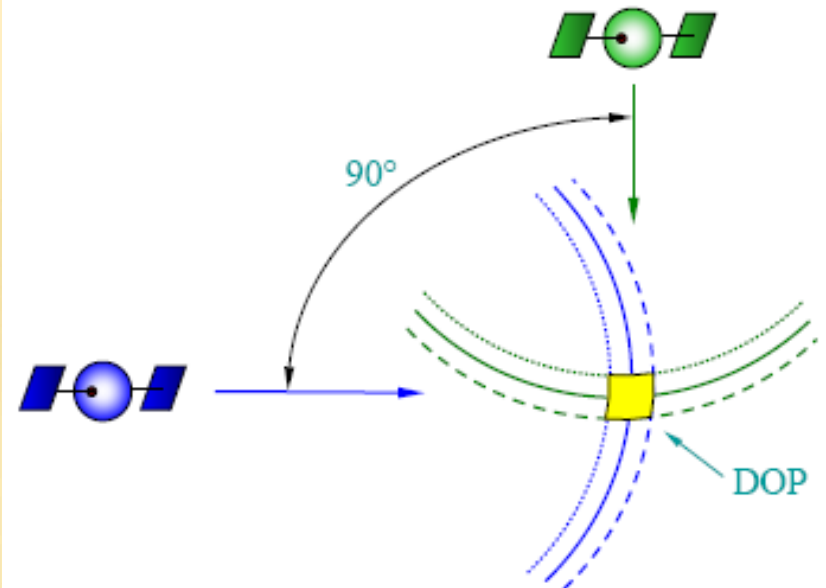
Can have problems similar to earthquake locating (two satellites in “same” direction for example – has effect of reducing rank by one)

GPS tutorial Signals and Data

Good DOP

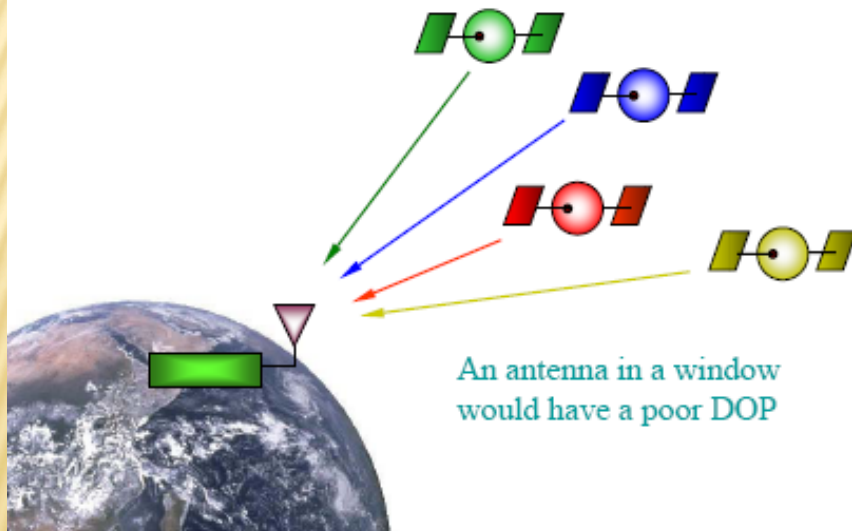


Intersecting Ranges

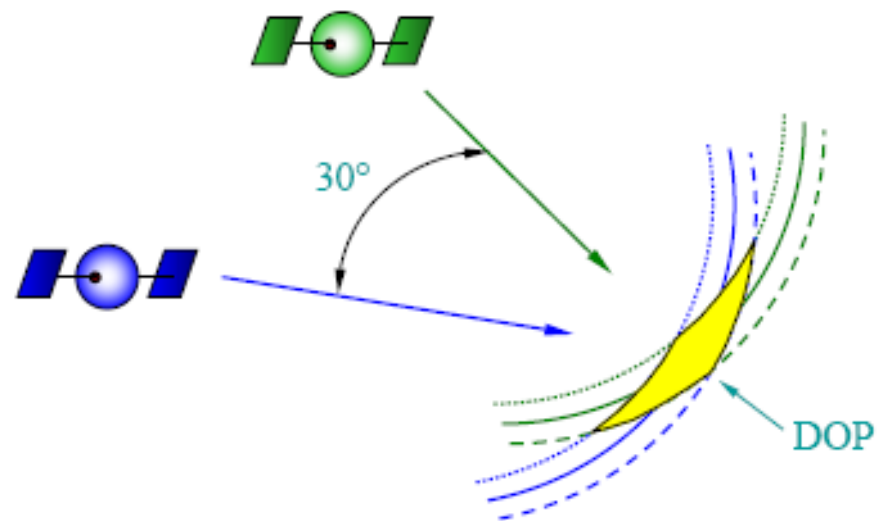


GPS tutorial Signals and Data

Poor DOP



Intersecting Ranges



Elementary Concepts

Variables:

things that we measure, control, or manipulate in research. They differ in many respects, most notably in the role they are given in our research and in the type of measures that can be applied to them.

Observational vs. experimental research.

Most empirical research belongs clearly to one of those two general categories.

In observational research we do not (or at least try not to) influence any variables but only measure them and look for relations (correlations) between some set of variables.

In experimental research, we manipulate some variables and then measure the effects of this manipulation on other variables.

Observational vs. experimental research.

Dependent vs. independent variables.

Independent variables are those that are manipulated
whereas dependent variables are only measured or
registered.

Variable Types and Information Content

Measurement scales.

Variables differ in "how well" they can be measured.

Measurement error involved in every measurement, which determines the "amount of information" obtained.

Another factor is the variable's "type of measurement scale."

Variable Types and Information Content

Nominal variables

allow for only qualitative classification.

That is, they can be measured only in terms of whether the individual items belong to some distinctively different categories, but we cannot quantify or even rank order those categories.

Typical examples of nominal variables are gender, race, color, city, etc.

Variable Types and Information Content

Ordinal variables

allow us to rank order the items we measure

in terms of which has less and which has more of the quality represented by the variable, but still they do not allow us to say "how much more."

A typical example of an ordinal variable is the socioeconomic status of families.

Variable Types and Information Content

Interval variables

allow us not only to rank order the items that are measured,

but also to quantify and compare the sizes of differences between them.

For example, temperature, as measured in degrees Fahrenheit or Celsius, constitutes an interval scale.

Variable Types and Information Content

Ratio variables

are very similar to interval variables;

in addition to all the properties of interval variables, they feature an identifiable absolute zero point, thus they allow for statements such as x is two times more than y .

Typical examples of ratio scales are measures of time or space.

Systematic and Random Errors

Error:

Defined as the difference between a calculated or observed value and the “true” value

Systematic and Random Errors

Blunders:

Usually apparent

either as obviously incorrect data points or results that are not reasonably close to the expected value.

Easy to detect (usually).

Easy to fix (throw out data).

Systematic and Random Errors

Systematic Errors:

Errors that occur reproducibly from faulty calibration of equipment or observer bias.

Statistical analysis is generally not useful, but rather corrections must be made based on experimental conditions.

Systematic and Random Errors

Random Errors:

Errors that result from the fluctuations in observations.

Requires that experiments be repeated a sufficient number of time to establish the precision of measurement.

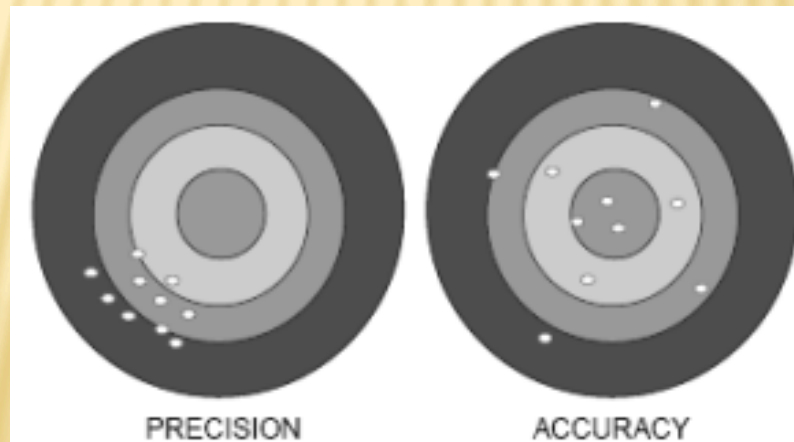
(statistics useful here)

Accuracy vs. Precision

Accuracy vs. Precision

Accuracy: A measure of how close an experimental result is to the true value.

Precision: A measure of how exactly the result is determined. It is also a measure of how reproducible the result is.



Accuracy vs. Precision

Absolute precision:

indicates the uncertainty in the same units as the observation

Relative precision:

indicates the uncertainty in terms of a fraction of the value of the result

Uncertainties

In most cases,

cannot know what the “true” value is unless there is an independent determination

(*i.e.* different measurement technique).

Uncertainties

Only can consider estimates of the error.

Discrepancy is the difference between two or more observations. This gives rise to uncertainty.

Probable Error:

Indicates the magnitude of the error we estimate to have made in the measurements.

Means that if we make a measurement that we will be wrong by that amount “on average”.

Parent vs. Sample Populations

Parent population:

Hypothetical probability distribution if we were to make an infinite number of measurements of some variable or set of variables.

Parent vs. Sample Populations

Sample population:

Actual set of experimental observations or measurements of some variable or set of variables.

In General:

(Parent Parameter) $\approx$ limit (Sample Parameter)

When the number of observations, N , goes to infinity.

some univariate statistical terms:

mode:

value that occurs most frequently in a distribution

(usually the highest point of curve)

may have more than one mode
(eg. Bimodal – example later)
in a dataset

some univariate statistical terms:

median:

value midway in the frequency distribution
...half the area under the curve is to right and other to left

mean:

arithmetic average
...sum of all observations divided by # of observations

the mean is a poor measure of central tendency in skewed distributions

Average, mean or expected value for random variable

$$E(x) = \mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i$$

(more general) if have probability for each x_i

$$E(x) = \mu = \lim_{n \rightarrow \infty} \sum_{i=1}^n p_i x_i$$

some univariate statistical terms:

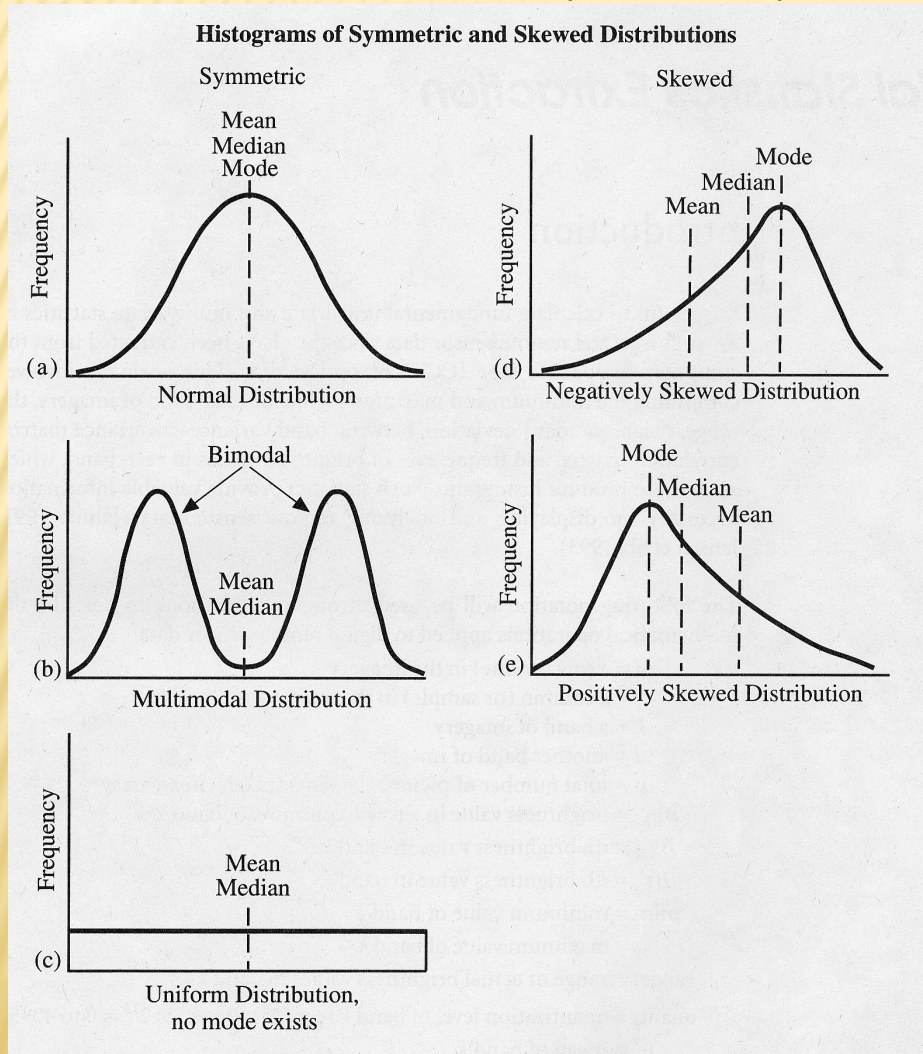
range: measure of dispersion about mean

(maximum minus minimum)

when max and min are unusual values, range may be a misleading measure of dispersion

Histogram

useful graphic representation of information content of sample or parent population

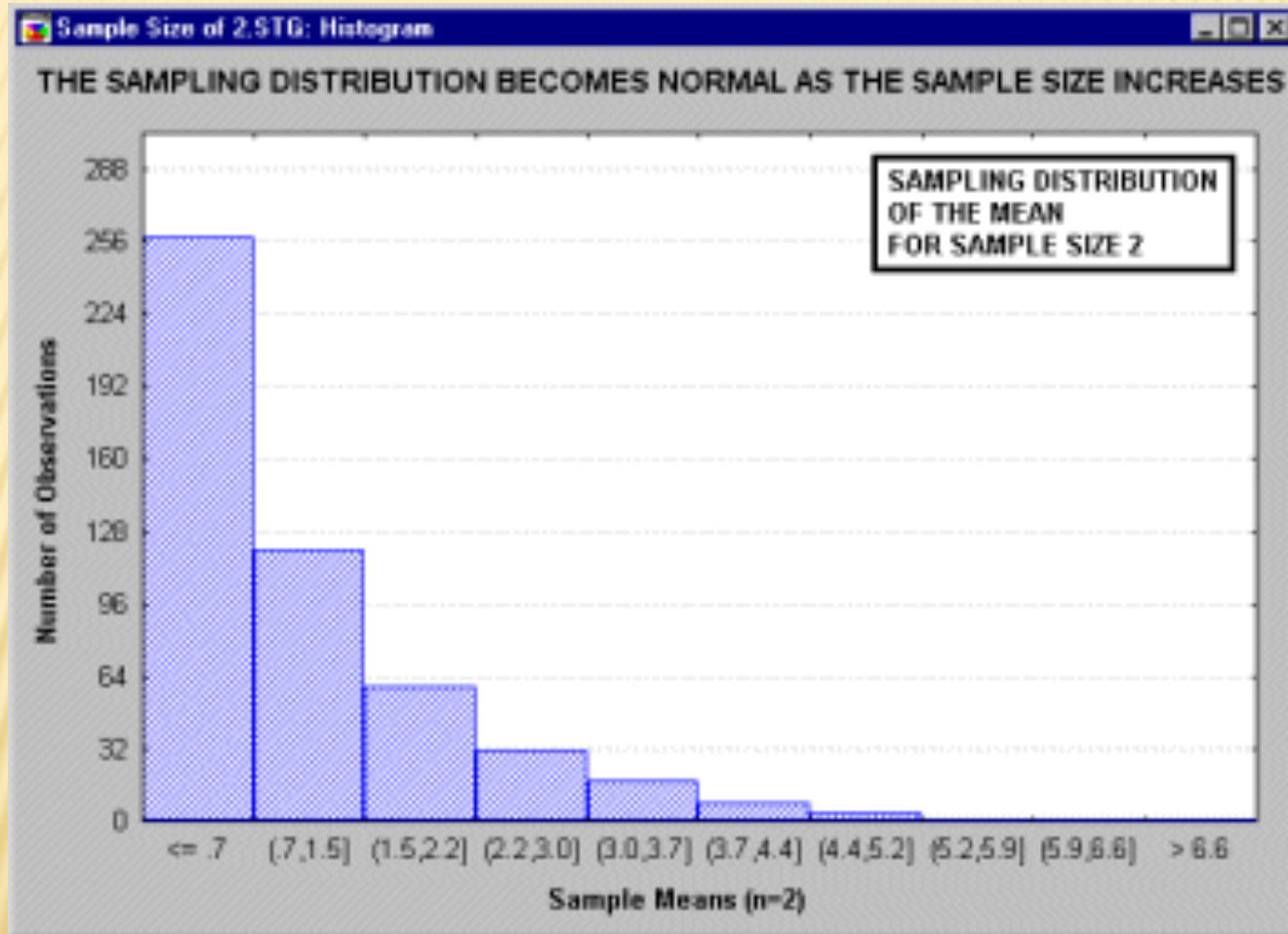


many statistical tests
assume
values are normally
distributed

*not always the case!
examine data prior
to processing*

from: Jensen, 1996

Distribution vs. Sample Size



Deviations

The deviation, δ_i , of any measurement x_i from the mean μ of the parent distribution is defined as the difference between x_i and μ

$$\delta x_i = x_i - \mu$$

Deviations

Average deviation, α ,

is defined as the average of the magnitudes
of the deviations,

Magnitudes given by the absolute value of the
deviations.

$$\alpha = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |x_i - \mu|$$

Root mean square

$$RMS = \lim_{n \rightarrow \infty} \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

Of deviations or residuals – standard deviation

$$\sigma = \lim_{n \rightarrow \infty} \sqrt{\frac{1}{n} \sum_{i=1}^n \delta_i^2}$$

Sample Mean and Standard Deviation

For a series of n observations, the most probable estimate of the mean μ is the average of the observations.

We refer to this as the sample mean to distinguish it from the parent mean μ .

$$\mu \approx \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Sample Mean and Standard Deviation

Our best estimate of the standard deviation σ would be from:

$$\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \approx \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

But we cannot know the true parent mean μ so the best estimate of the sample variance and standard deviation would be:

$$\sigma^2 = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{Sample Variance}$$

Some other forms to write variance

$$\sigma^2 = \text{VAR}(x) = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - E(x))^2 = \frac{1}{(n-1)} \left(\sum_{i=1}^n x_i^2 - N\bar{x} \right)$$

$$\delta x_i = x_i - \bar{x} \qquad \sigma^2 = \frac{1}{(n-1)} \sum_{i=1}^n \delta x_i^2$$

If have probability for each x_i

$$\sigma^2 = \text{VAR}(x) = \sum_{i=1}^n p_i (x_i - E(x))^2$$

The standard deviation

$$\sigma = \sqrt{\text{VAR}(x)} = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n \delta x_i^2}$$

(Normalization decreased from N to $(N-1)$ for the “sample” variance, as μ is used in the calculation)

For a scalar random variable or measurement with a Normal (Gaussian) distribution,

the probability of being within one σ of the mean is 68.3%

small std dev:

observations are clustered tightly about the mean

large std dev:

observations are scattered widely about the mean

Distributions

Binomial Distribution: Allows us to define the probability, p , of observing x a specific combination of n items, which is derived from the fundamental formulas for the permutations and combinations.

Permutations: Enumerate the number of permutations, $P_m(n, x)$, of coin flips, when we pick up the coins one at a time from a collection of n coins and put x of them into the “heads” box.

$$P_m(n, x) = \frac{n!}{(n - x)!}$$

Combinations:

Relates to the number of ways we can combine the various permutations enumerated above from our coin flip experiment.

Thus the number of combinations is equal to the number of permutations divided by the degeneracy factor $x!$ of the permutations (number indistinguishable permutations) .

$$C(n, x) = \frac{P_m(n, x)}{x!} = \frac{n!}{x!(n - x)!} = \binom{n}{x}$$

Probability and the Binomial Distribution

Coin Toss Experiment: If p is the probability of success
(landing heads up)
is not necessarily equal to the probability $q = 1 - p$ for
failure
(landing tails up) because the coins may be lopsided!

The probability for each of the combinations of x coins
heads up and
 $n - x$ coins tails up is equal to $p^x q^{n-x}$.

The binomial distribution can be used to calculate the
probability:

Probability and the Binomial Distribution

The binomial distribution can be used to calculate the probability of x “successes” in n tries where the individual probability is p :

$$P_B(n, x, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

The coefficients $P_B(x, n, p)$ are closely related to the binomial theorem for the expansion of a power of a sum

$$(p + q)^n = \sum_{x=0}^n \binom{n}{x} p^x q^{n-x}$$

Mean and Variance: Binomial Distribution

$$\mu = \left[\sum_{x=0}^n x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np$$

The average of the number of successes will approach a mean value μ

given by the probability for success of each item p times the number of items.

For the coin toss experiment $p=1/2$, half the coins should land heads up on average.

Mean and Variance: Binomial Distribution

The standard deviation is

$$\sigma^2 = \sum_{x=0}^n \left[(x - \mu)^2 \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np(1-p)$$

If the the probability for a single success p is equal to the probability for failure $p=q=1/2$,

the final distribution is symmetric about the mean,
and mode and median equal the mean.

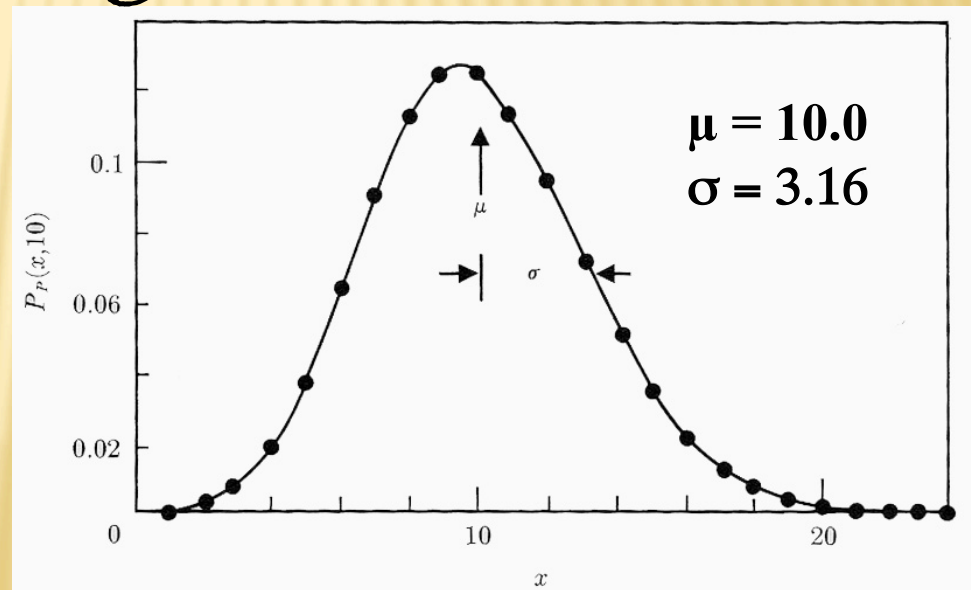
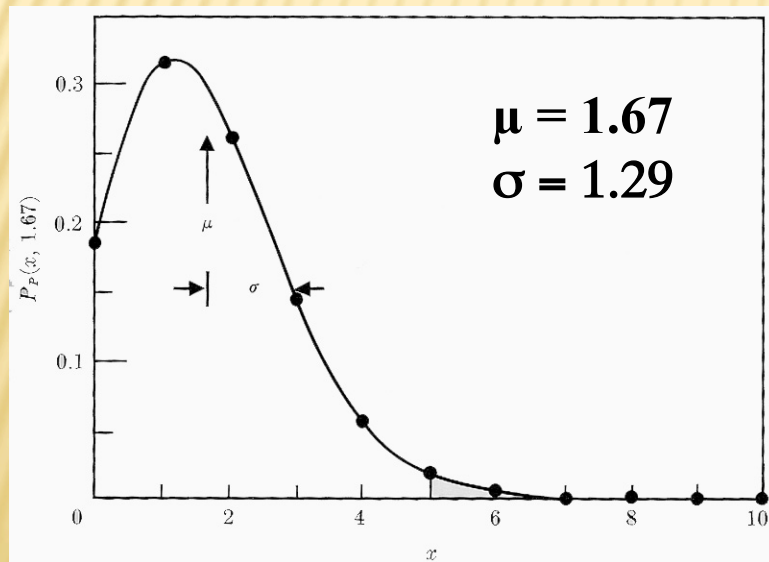
The variance, $\sigma^2 \approx m/2$.

Other Probability Distributions: Special Cases

Poisson Distribution: Approximation to binomial distribution for special case when average number of successes is very much smaller than possible number *i.e.*
 $\mu \ll n$ because $p \ll 1$.

Distribution is NOT necessarily symmetric! Data are usually bounded on one side and not the other.

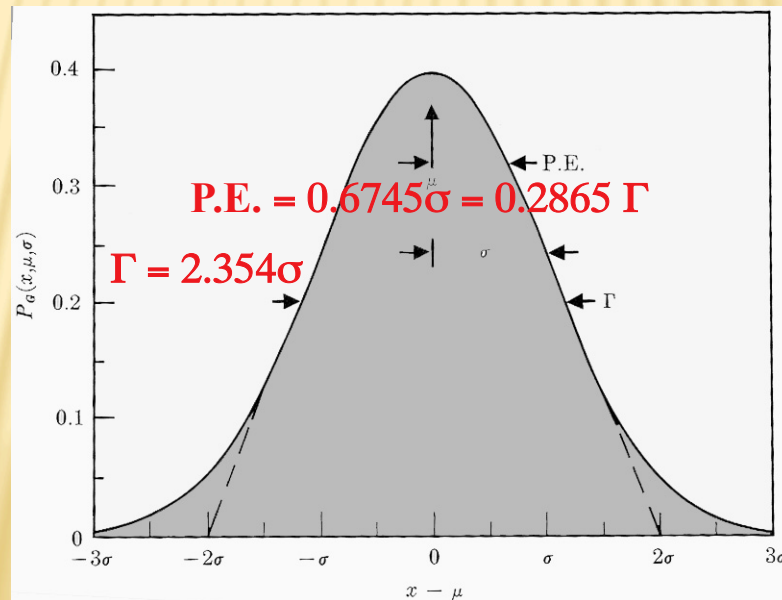
Advantage $\sigma^2 \approx \mu$.



Gaussian or Normal Error Distribution

Gaussian Distribution: Most important probability distribution in the statistical analysis of experimental data.

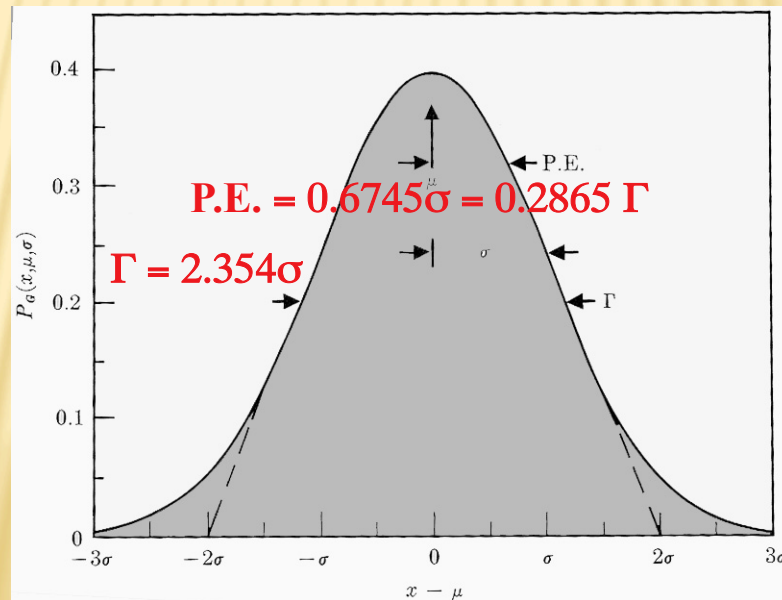
Functional form is relatively simple and the resultant distribution is reasonable.



Gaussian or Normal Error Distribution

Another special limiting case of binomial distribution where the number of possible different observations, n , becomes infinitely large yielding $np \gg 1$.

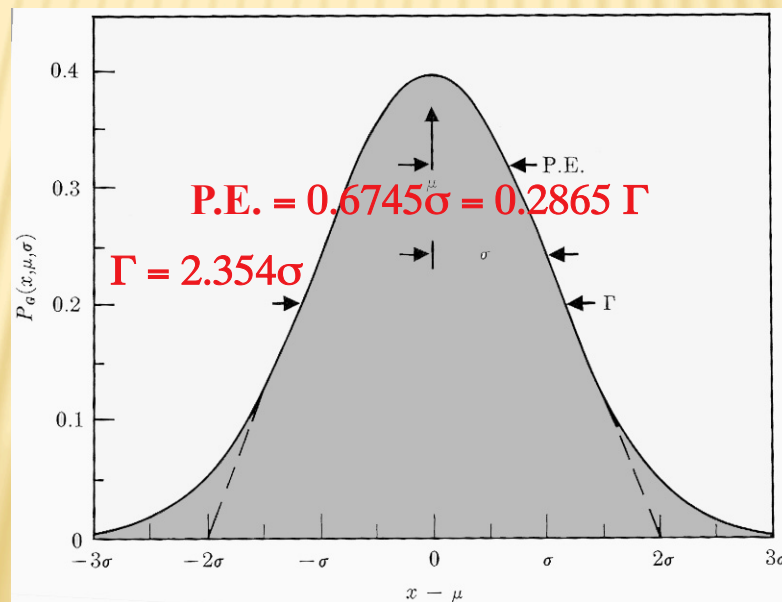
Most probable estimate of the mean μ from a random sample of observations is the average of those observations!



Gaussian or Normal Error Distribution

Probable Error (P.E.) is defined as the absolute value of the deviation such that P_G of the deviation of any random observation is $< \frac{1}{2}$

Tangent along the steepest portion of the probability curve intersects at $e^{-1/2}$ and intersects x axis at the points $x \approx \mu \pm 2\sigma$



For gaussian / normal error distributions:

Total area underneath curve is 1.00 (100%)

68.27% of observations lie within ± 1 std dev of mean

95% of observations lie within ± 2 std dev of mean

99% of observations lie within ± 3 std dev of mean

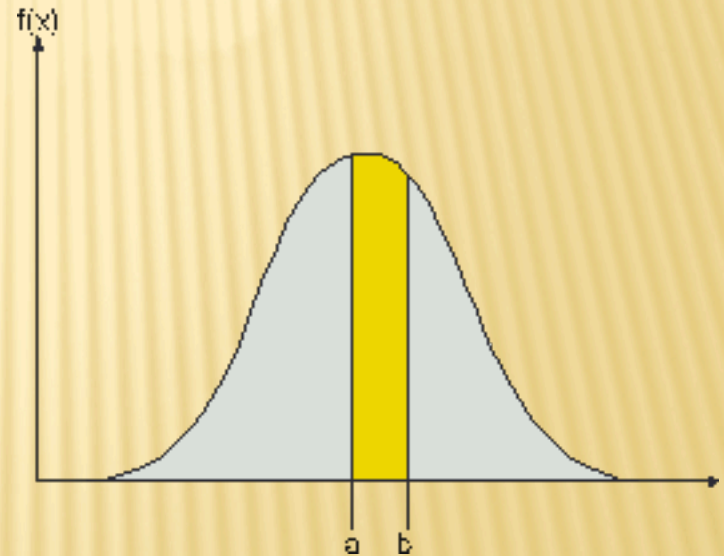
Variance, standard deviation, probable error, mean, and weighted root mean square error are commonly used statistical terms in geodesy.

compare (rather than attach significance to numerical value)

If X is a continuous random variable, then the probability density function, pdf, of X ,

is a function $f(x)$ such that for two numbers, a and b with $a \leq b$

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$



That is, the probability that X takes on a value in the interval $[a, b]$ is the area under the density function from a to b .

The probability density function for the Gaussian distribution is defined as:

$$P_G(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right]$$

For the Gaussian PDF, the probability for the random variable x to be found between $\mu \pm z\sigma$,
Where z is the dimensionless range $z = |x - \mu|/\sigma$ is:

$$A_G(x, \mu, \sigma) = \int_{\mu - z\sigma}^{\mu + z\sigma} P_G(x, \mu, \sigma) dx = \frac{1}{\sqrt{2\pi}} \int_{-z}^z \exp\left[-\frac{1}{2} x^2\right] dx$$

$$A_G(z = \infty) = 1$$

The cumulative distribution function, cdf,
is a function $F(x)$ of a random variable, X , and
is defined for a number x by:

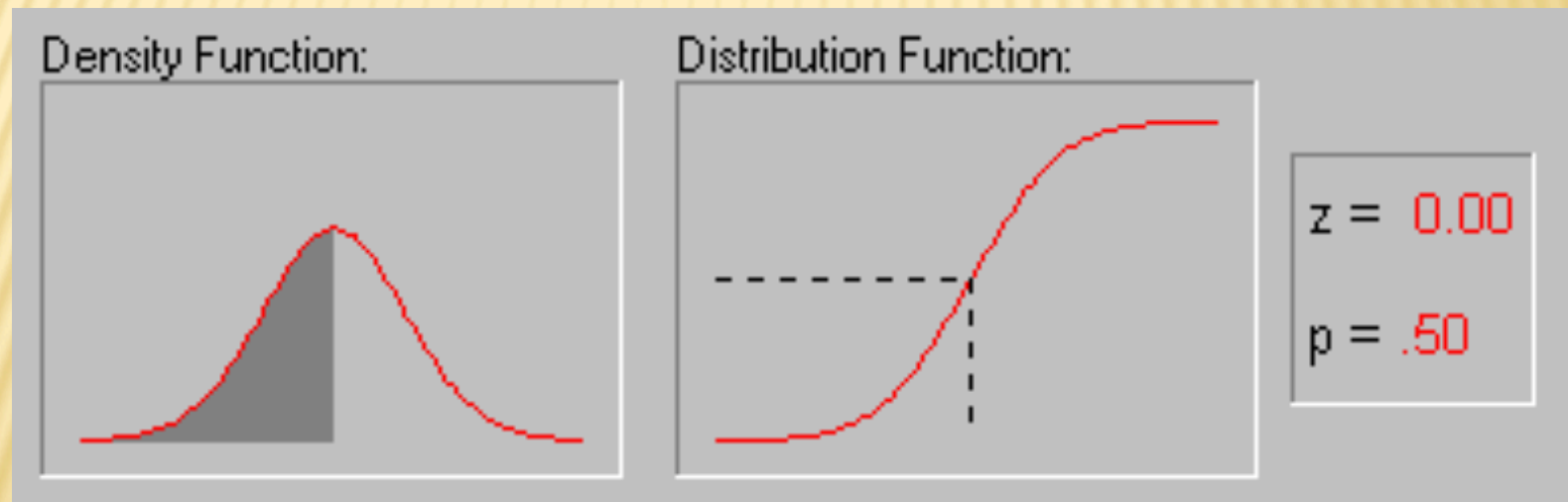
$$F(x) = P(X \leq x) = \int_{0,\infty}^x f(s)ds$$

That is, for a given value x , $F(x)$ is the probability that the observed value of X will be at most x . (note lower limit shows domain of s , integral goes from 0 to $x < \infty$)

Relationship between PDF and CDF

Density vs. Distribution Functions for Gaussian

<~ derivative <~



-> integral ->

Multiple random variables

Expected value or mean of sum of two random variables
is sum of the means.

known as additive law of expectation.

$$E(x + y) = E(x) + E(y)$$

covariance

$$\sigma_{xy}^2 = COV(x, y) = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - E(x))(y_i - E(y))$$

(variance is covariance of variable with itself)

(more general with) individual probabilities

$$\sigma_{xy}^2 = COV(x, y) = \sum_{i=1}^n p_{xy_i} (x_i - E(x))(y_i - E(y))$$

Covariance matrix

$$\delta x_i = x_i - \bar{x}$$

$$\delta y_i = y_i - \bar{y}$$

$$C = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{(n-1)} \sum_{i=1}^n \delta x_i^2 & \frac{1}{(n-1)} \sum_{i=1}^n \delta x_i \delta y_i \\ \frac{1}{(n-1)} \sum_{i=1}^n \delta x_i \delta y_i & \frac{1}{(n-1)} \sum_{i=1}^n \delta y_i^2 \end{pmatrix}$$

Covariance matrix defines error ellipse.

Eigenvalues are squares of semimajor and semiminor axes (σ_1 and σ_2)

Eigenvectors give orientation of error ellipse

(or given σ_x and σ_y , correlation gives “fatness” and “angle”)

Distance Root Mean Square (DRMS, 2-D extension of RMS)

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad \Rightarrow$$

$$DRMS = \left(\sigma_x^2 + \sigma_y^2 \right)^{\frac{1}{2}}$$

For a scalar random variable or measurement with a Normal (Gaussian) distribution,
the probability of being within the 1- σ ellipse about the mean is 68.3%

Etc for 3-D

Use of variance, covariance – in Weighted Least Squares

common practice to use the reciprocal of the variance as
the weight

variance of the sum of two random variables

$$VAR(x + y) = VAR(x) + 2COV(x, y) + VAR(y)$$

The variance of the sum of two random variables is equal to the sum of each of their variances only when the random variables are independent

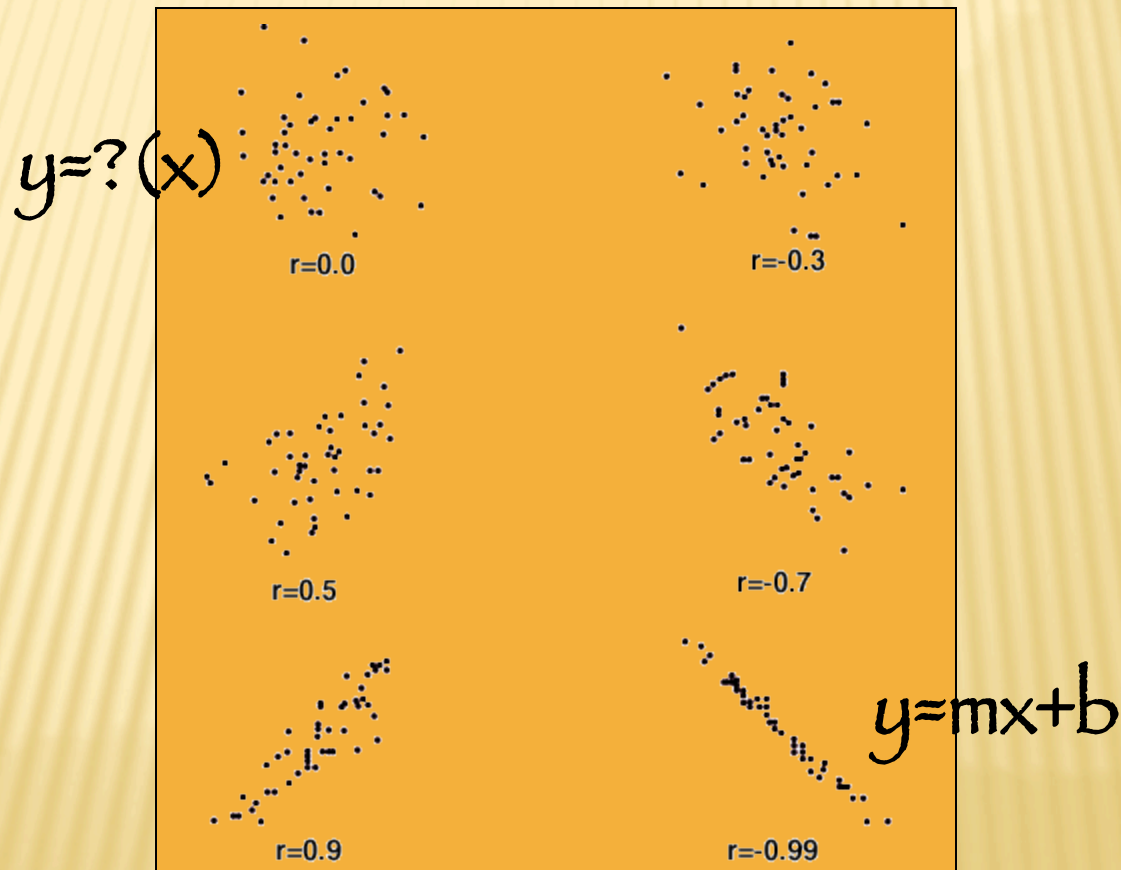
(The covariance of two independent random variables is zero, $cov(x, y) = 0$).

Multiplying a random variable by a constant increases the variance by the square of the constant.

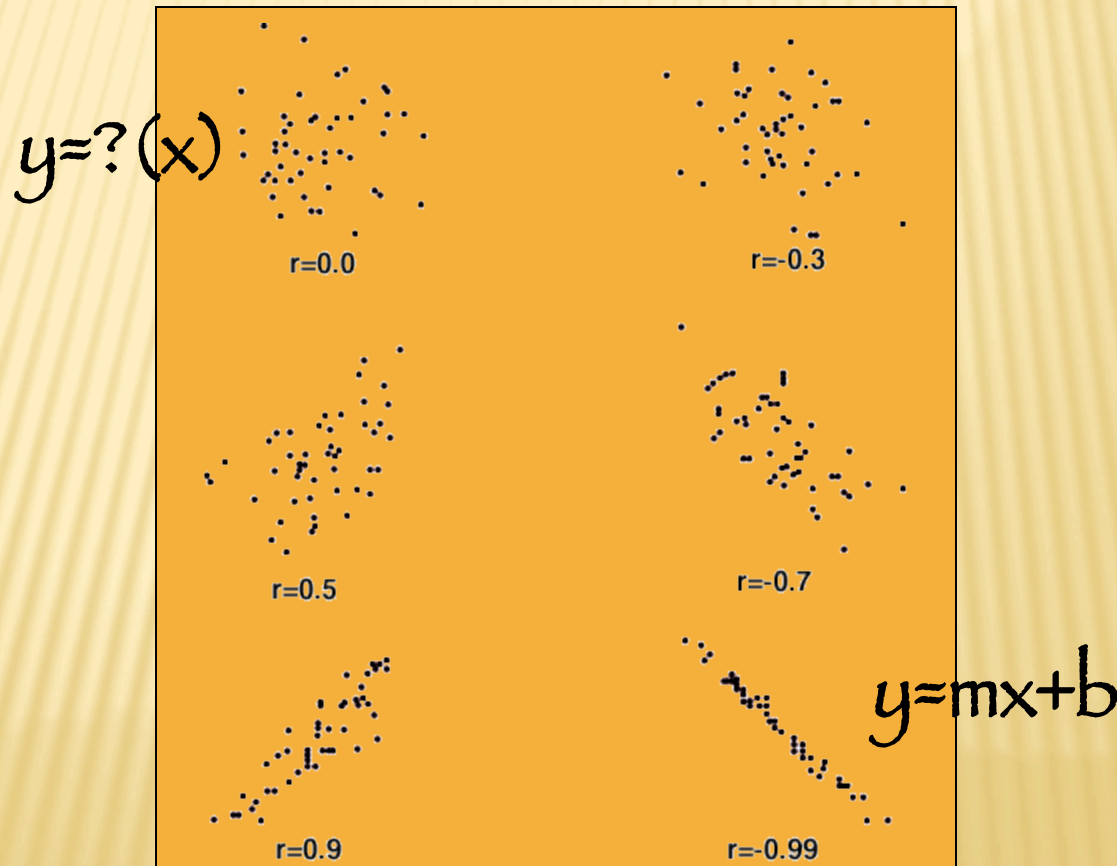
$$\sigma_{cx}^2 = E(cx) = c^2 E(x)$$

Correlation

The more tightly the points are clustered together the higher the correlation between the two variables and the higher the ability to predict one variable from another

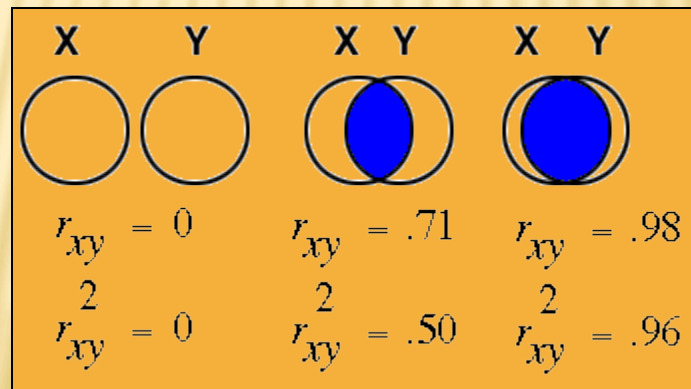


Correlation coefficients are between -1 and +1,
+ and - 1 represent perfect correlations,
and zero representing no relationship, between the
variables.

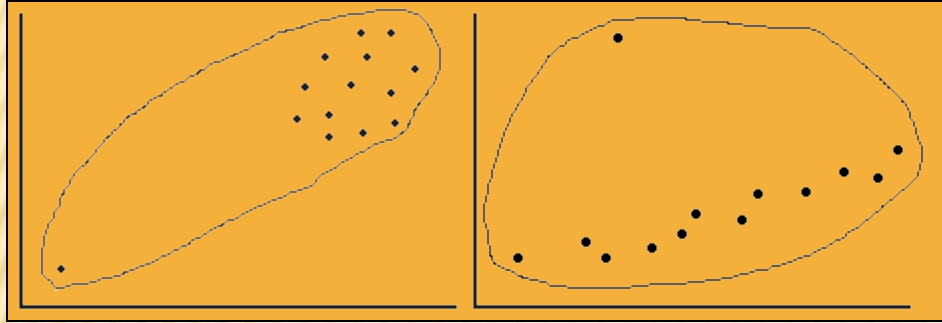


Correlations are interpreted by squaring the value of the correlation coefficient.

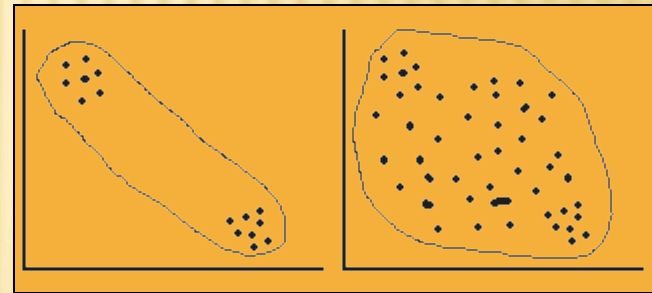
The squared value represents the proportion of variance of one variable that is shared with the other variable, in other words, the proportion of the variance of one variable that can be predicted from the other variable.



Sources of misleading correlation (and problems with least squares inversion)

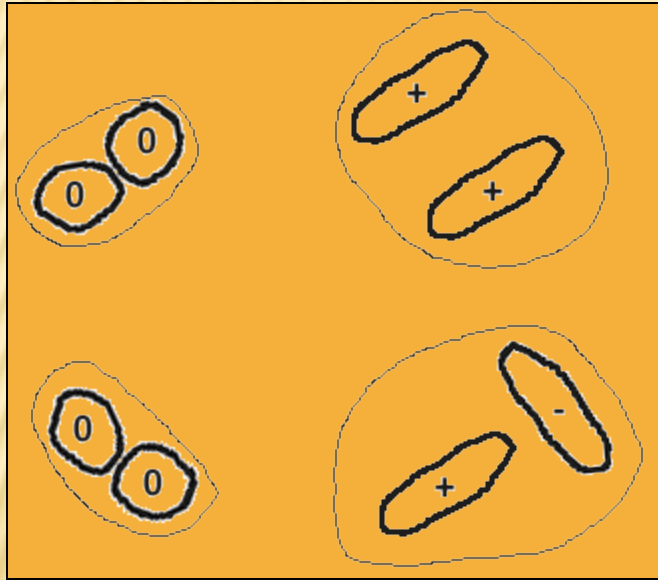


outliers

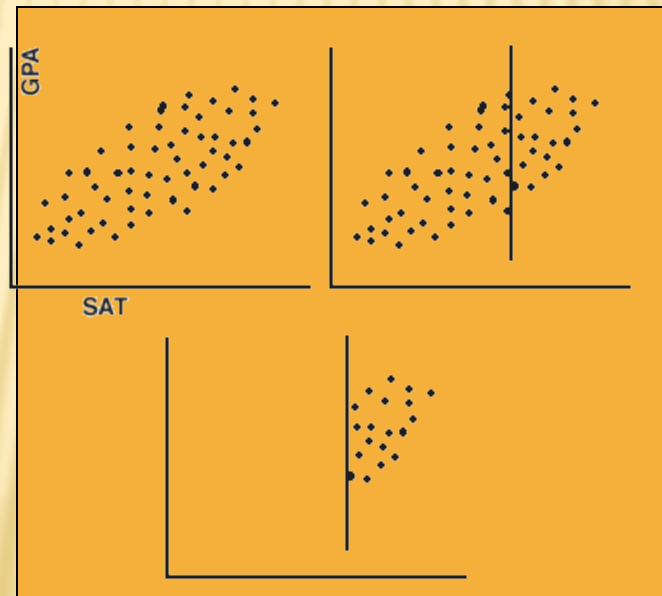


Bimodal
distribution relation

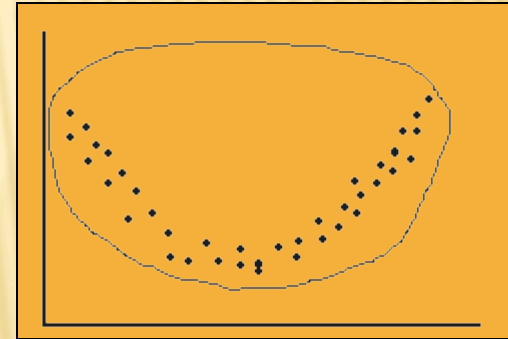
Sources of misleading correlation (and problems with least squares inversion)



Combining
groups



Restriction of range



curvilinearity

rule of thumb for interpreting correlation coefficients:

Corr	Interpretation
0 to .1	trivial
.1 to .3	small
.3 to .5	moderate
.5 to .7	large
.7 to .9	very large

Correlations express the inter-dependence between variables.

For two variables x and y in a linear relationship, the correlation between them is defined as

$$\chi_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

High correlation does not mean that the variations of one are caused by the variations of the others, although it may be the case.

In many cases, external influences may be affecting both variables in a similar fashion.

two types of correlation

physical correlation and mathematical correlation

Physical correlation refers to the correlations between the actual field observations.

It arises from the nature of the observations as well as their method of collection.

If different observations or sets of observation are affected by common external influences, they are said to be physically correlated.

Hence all observations made at the same time at a site may be considered physically correlated because similar atmospheric conditions and clock errors influence the measurements.

Mathematical correlation is related to the parameters in the mathematical model.

It can therefore be partitioned into two further classes which correspond to the two components of the mathematical adjustment model:

Functional correlation

Stochastic correlation

Functional Correlation:

The physical correlations can be taken into account by introducing appropriate terms into the *functional model* of the observations.

That is, functionally correlated quantities share the same parameter in the observation model.

An example is the clock error parameter in the one-way GPS observation model, used to account for the physical correlation introduced into the measurements by the receiver clock and/or satellite clock errors.

Stochastic Correlation:

Stochastic correlation (or statistical correlation) occurs between observations when non-zero off-diagonal elements are present in the variance-covariance (VCV) matrix of the observations.

Also appears when functions of the observations are considered (eg. differencing), due to the Law of Propagation of Variances.

However, even if the VCV matrix of the observations is diagonal (no stochastic correlation), the VCV matrix of the resultant LS estimates of the parameters will generally be full matrices, and therefore exhibit stochastic correlation.