









Usefulness: Correlation (r)

- Correlation is a measure of how much two sets of data are related. CAUTION: "Correlation does not imply causation.
- r is the symbol for correlation. r varies from -1.0 to +1.0.
- +1.0 indicates a perfect positive relationship, -1.0 indicates a perfect inverse relationship, 0 indicates no relationship.
- Ex: If age and height were perfectly related (r=1.0), knowing one's age would allow exact knowledge of one's height, with height increasing as age increases. However, since age and height are not perfectly related, knowing age allows prediction of height with an associated amount variation.
- If time spent on technical reviews were inversely related to number of defects reported, with, say r=-0.75, then knowing the amount of time spent on reviews would allow us to predict relatively accurately the number of defects that would be reported in the delivered AU INSY 560, Singapore 1997, Dan Turk





Usefulness: Significance (a)

- A correlation's a (alpha) indicates the statistical significance of the correlation, or the degree to which the correlation would be expected simply due to chance. The smaller the a the less likely the results are simply due
- to chance. 0.05 is generally accepted as a "good" a level; sometimes even up to 0.1 is accepted.
- Ex: If data are collected 100 times and we determine r=0.75 with a=0.05, then we would expect to obtain a correlation of 0.75 5 times simply because of random chance. If r=0.75 occurs more than 5 times in 100 then we might conclude that there is something other than random fluctuation which is causing it to occur. AU INSY 560. Singapore 1997. Dan Turk Humphrey Ch. 4 - slide 9





AU INSY 560, Singapore 1997, Dan Turk

ev 1995 lecture slide Imprecise and inaccurate Precise and inaccurate mprecise and accurate Precise and accurate

Precision vs. Accuracy

Note: Humphrey is using the term "precision" here in the sense of "reliability", that of obtaining the same results over and over, not in the sense of "granularity" AU INSY 560, Singapore 1997, Dan Turk Humphrey Ch. 4 - slide 12



Accuracy requires a relevant collection of historical data and a method for extrapolating from it.

AU INSY 560, Singapore 1997, Dan Turk









	Counting Example (based on Humphrey, 1995, lecture notes example and Dan Turk's counting standard)
	<pre>// Lset_set() determines if n is in an "integer set" collection. // It returns TRUE if so, and FALSE if not. void iset_set (</pre>
	<pre>inc = FALSE; search_ptr = set_start; while (search_ptr != NULL && inc == FALSE) { if (searth_ptr-sthis_n == *n) inc = TRUE; else search_ptr = search_ptr->next_n;</pre>
	} // wime }// iset_set()
AU INSY :	// In this example, LOC = 12. 560, Singapore 1997, Dan Turk Humphrey Ch. 4 - slide 19







Using LOC: Evaluating Development Work (cf. Humphrey, 1995, p. 81-82)

- Use LOC count numbers which seem most appropriate to the task at hand. The correct answer will vary from task to task.
- Ex: When evaluating development status (where are we?), new and changed LOC is probably most relevant.
- Ex: When forecasting a new product, don't consider unmodified reused code.

AU INSY 560, Singapore 1997, Dan Turk



AU INSY 560, Singapore 1997, Dan Turk









