Intermediate Laboratory – PHX 3870
Lecture Five

Chauvenet's Criterion

Consider the following example of the application of Chauvenet's Criterion to determine if a certain datum should be rejected.

A set of N=10 measurements of a length are made. The data are assumed to be described by a randon Gaussian distribution.

<u>Enter Data</u>

Number of data points: $N := 10$

Data indices: $i := 0 .. (N - 1)$

Enter data set: $x_i :=$

$45.7 \cdot m$
$46.2 \cdot m$
$46.9 \cdot m$
$54.8 \cdot m$
$46.1 \cdot m$
$45.2 \cdot m$
$45.4 \cdot m$
$47.0 \cdot m$
$45.9 \cdot m$
$46.3 \cdot m$

Calculate mean: $x_{mean} := \mathrm{mean}(x) = 46.95\,m$

Calculate standard deviation: $\sigma_x := \mathrm{stdev}(x) = 2.673\,m$

Calculate fractional deviation from the mean: $Frac\_Dev_i := \left| \dfrac{x_i - x_{mean}}{\sigma_x} \right|$

To apply Chauvenet's criterion, we first sort the data x in order of ascending values of the fractional deviation from the mean. The probability that a data point is likely to fall outside a given deviation is then calculated. We then determine how many data that should be eliminated based on Chauvenet's

Sort data in ascending order:

The probability that a data point is likely to fall outside a given deviation is:

$$\mathrm{Prob}(x_{test}, X, \sigma) := 1 - \int_{-|x_{test}|}^{|x_{test}|} \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\left[\frac{(x-X)^2}{2\cdot\sigma^2}\right]} dx$$

Apply Chauvenet's criterion and determine how many data points should be rejected:

$$\mathrm{Reject}(x, X, \sigma, N) := \mathrm{if}\left[(N \cdot \mathrm{Prob}(x, X, \sigma)) > 50 \cdot \%, \text{"Keep"}, \text{"Reject"}\right]$$

$$N_{reject} := \sum_{i=0}^{N-1} \mathrm{if}\left[\left(N \cdot \mathrm{Prob}\left(x_i, x_{mean}, \sigma_x\right)\right) > 50 \cdot \%, 0, 1\right] = 1$$

$x =$, $\mathrm{Frac\_Dev} =$, $\mathrm{Prob}(x_i, x_{mean}, \sigma_x) =$, $N \cdot \mathrm{Prob}(x_i, x_{mean}, \sigma_x) =$, $\mathrm{Reject}(x_i, x_{mean}, \sigma_x, N) =$

	x	Frac_Dev	Prob	N · Prob	Reject
0	45.7	0.468	0.68	6.8	"Keep"
1	46.2	0.281	0.61	6.105	"Keep"
2	46.9	0.019	0.507	5.075	"Keep"
3	54.8	2.936	$1.66\cdot10^{-3}$	0.017	"Reject"
4	46.1	0.318	0.625	6.247	"Keep"
5	45.2	0.655	0.744	7.436	"Keep"
6	45.4	0.58	0.719	7.19	"Keep"
7	47	0.019	0.493	4.925	"Keep"
8	45.9	0.393	0.653	6.528	"Keep"
9	46.3	0.243	0.596	5.961	"Keep"

Now recalculate the mean and standard deviation after rejecting N_{reject} data points.

Truncated data set indices and data array: $j := 0 .. N - 1 - N_{reject}$ $\quad X_{CH_j} := x_{order_j}$

The final analysis is:

	<u>Including all data points</u>	<u>Excluding the rejected data point</u>
Number data points:	$N = 10$	$N - N_{reject} = 9$
Mean:	$X_{mean} := \mathrm{mean}(x) = 46.95\ \mathrm{m}$	$X_{mean} := \mathrm{mean}(X_{CH}) = ▮$
Standard Deviation:	$\sigma_x := \mathrm{stdev}(x) = 2.673\ \mathrm{m}$	$\sigma_x := \mathrm{stdev}(X_{CH}) = ▮$