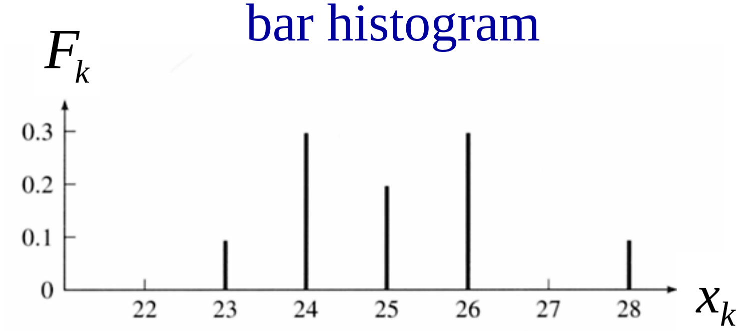
Histograms and Distributions

10 measurements:
26, 24, 26, 28, 23, 24, 25, 24, 26, 25

different values / number of occurrences						
different values x_k	23	24	25	26	27	28
number of occurrences n_k	1	3	2	3	0	1

bar histogram



$$F_k = \frac{n_k}{N}$$ fraction of measurements that gave the result x_k

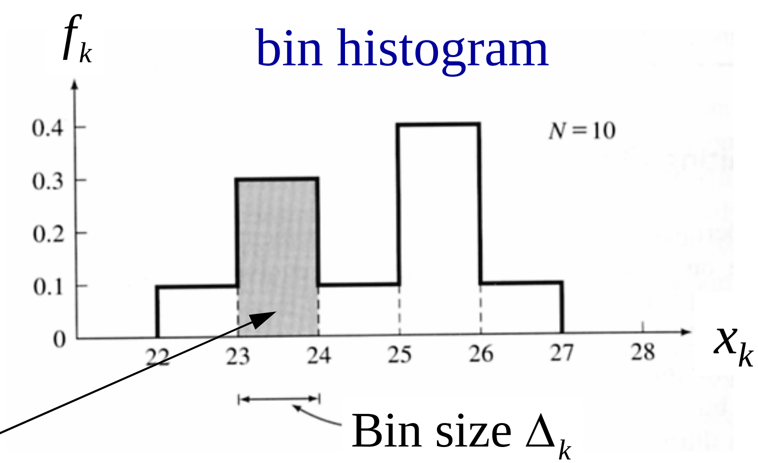
---

10 measurements:
26.4, 23.9, 25.1, 24.6, 22.7, 23.8, 25.1, 23.9, 25.3, 25.4

bins for x_k	22-23	23-24	24-25	25-26	26-27
n_k	1	3	1	4	1

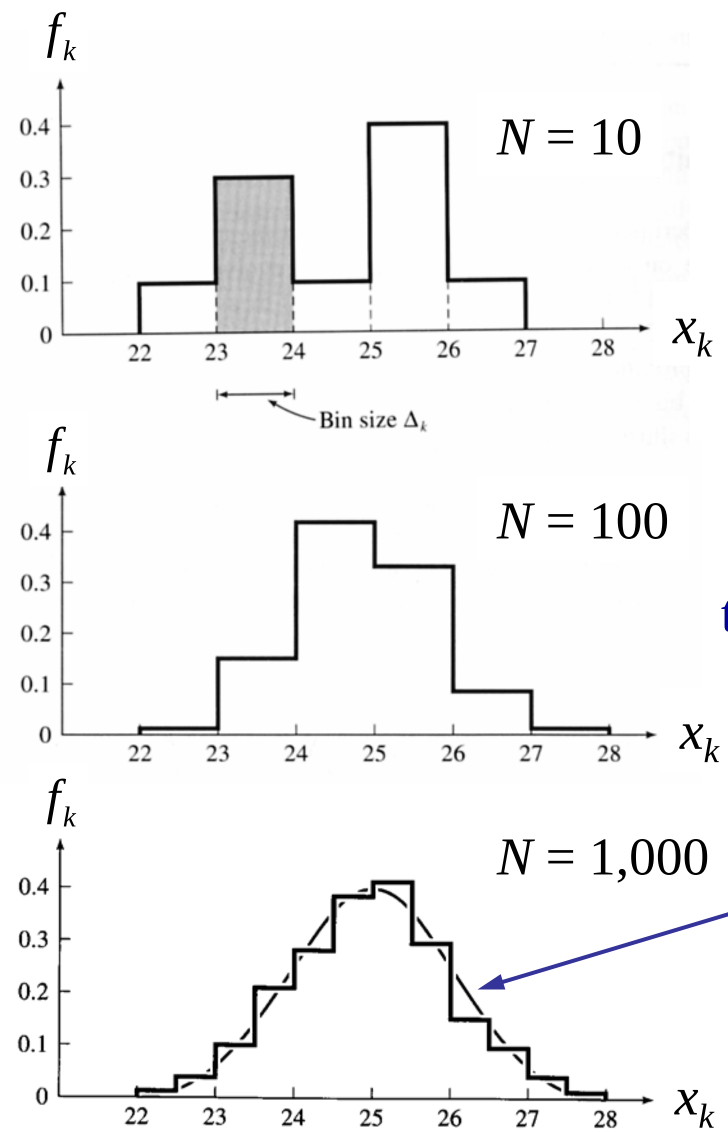
$$f_k \Delta_k = \frac{n_k}{N}$$ fraction of measurements in k-th bin

bin histogram



$f_k \Delta_k$ = the area of the k-th rectangle
has the same significance
as the height F_k of the k-th bar in a bar histogram

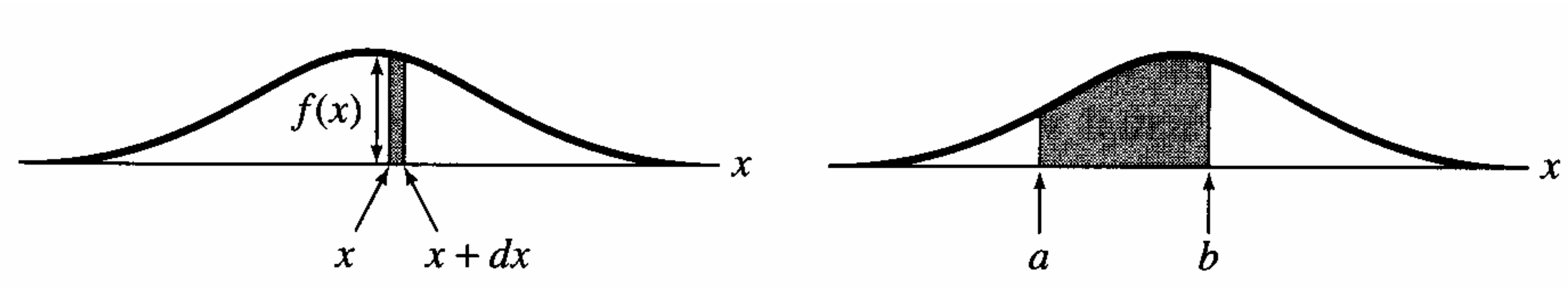
Limiting Distributions



as the number of measurements approaches infinity, their distribution approaches some definite continuous curve

this curve is called the limiting distribution, *f(x)*

Limiting Distributions

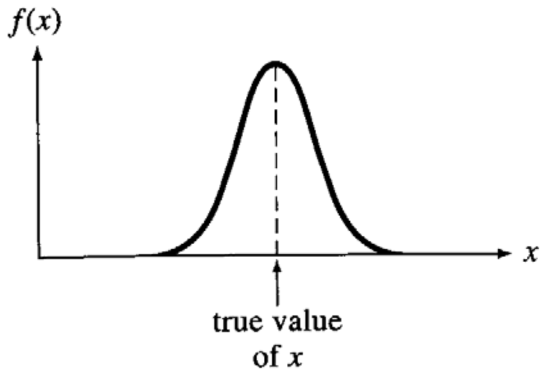


$f(x)\,dx$ = fraction of measurements that fall between x and $x+dx$
= probability that any measurement will give an answer between x and $x+dx$

$\int_a^b f(x)dx$ = fraction of measurements that fall between $x=a$ and $x=b$
= probability that any measurement will give an answer between $x=a$ and $x=b$

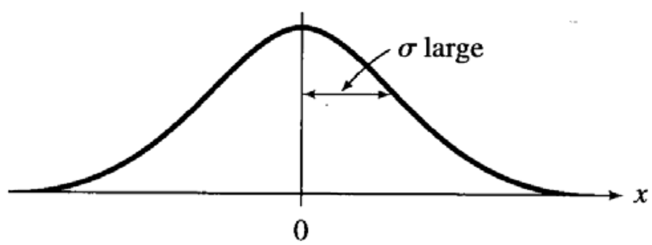
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$ normalization condition

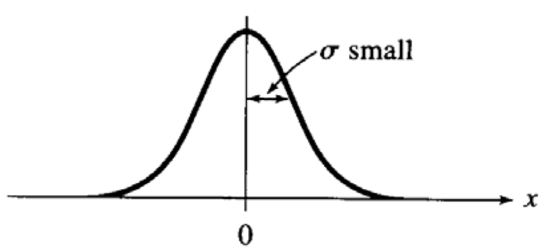
The Gauss, or Normal Distribution



the limiting distribution for a measurement subject to many small random errors is bell shaped and centered on the true value of x

the mathematical function that describes the bell-shape curve is called the normal distribution, or Gauss function





prototype function

$$e^{-x^2/2\sigma^2}$$

$$e^{-(x-X)^2/2\sigma^2}$$

σ – width parameter

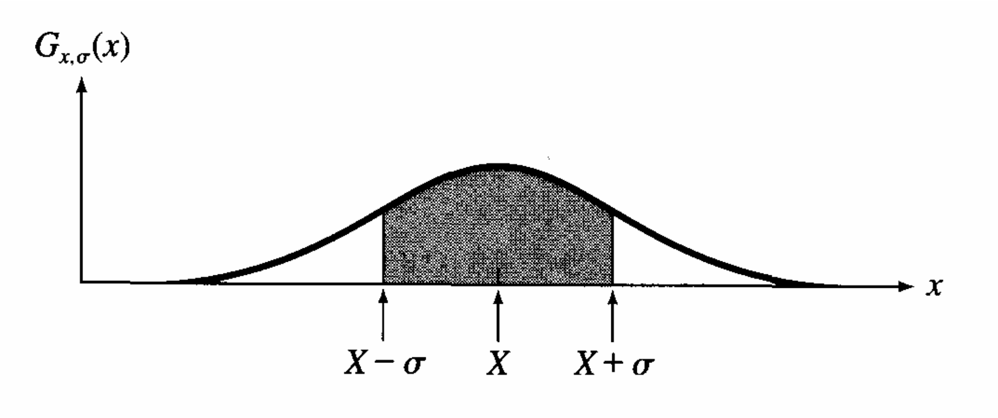
X – true value of x

The Gauss, or Normal Distribution

normalize $e^{-(x-X)^2/2\sigma^2}$ $\longrightarrow$ $\int_{-\infty}^{+\infty} f(x)dx = 1$

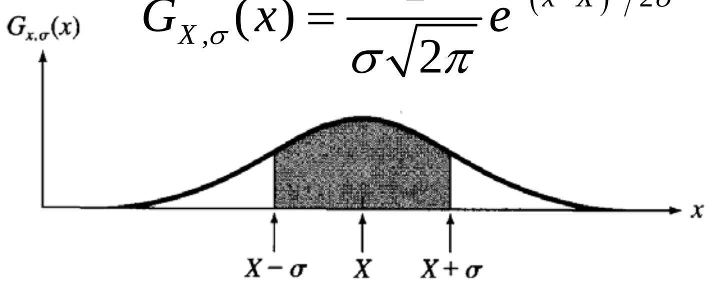
$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

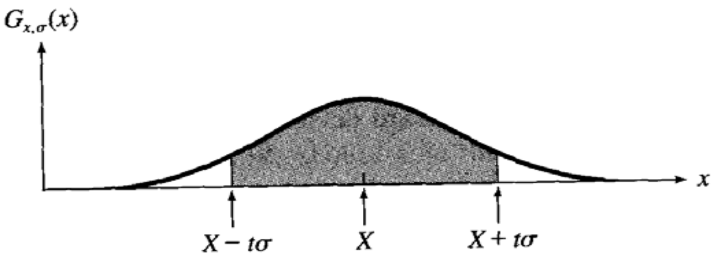
standard deviation σ_x = width parameter of the Gauss function σ
the mean value of x = true value X

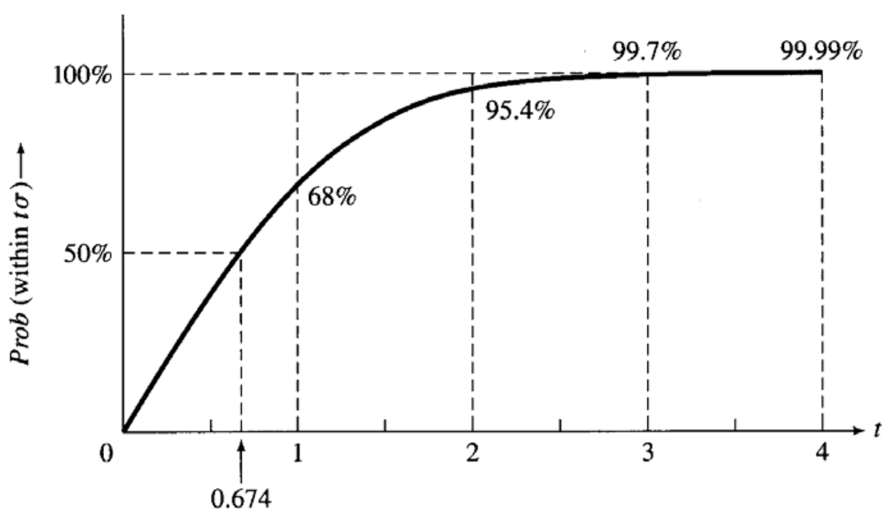


The standard Deviation as 68% Confidence Limit

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$







$$\text{Prob(within } \sigma) = \int_{X-\sigma}^{X+\sigma} G_{X,\sigma}(x)dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{X-\sigma}^{X+\sigma} e^{-(x-X)^2/2\sigma^2} dx$$

$$(x - X)/\sigma = z$$

$$\text{Prob(within } \sigma) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-z^2/2} dz$$

$$\text{Prob(within } t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-z^2/2} dz$$

erf(t) – error function

the probability that a measurement will fall within one standard deviation of the true answer is 68 %

$x = x_{best} \pm \delta x \qquad \delta x = \sigma$

Example:

A student measures a quantity x many times and calculates the mean as $\overline{x} = 10$ and the standard deviation as $\sigma_x = 1$. What fraction of his readings would you expect to find between 11 and 12?

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

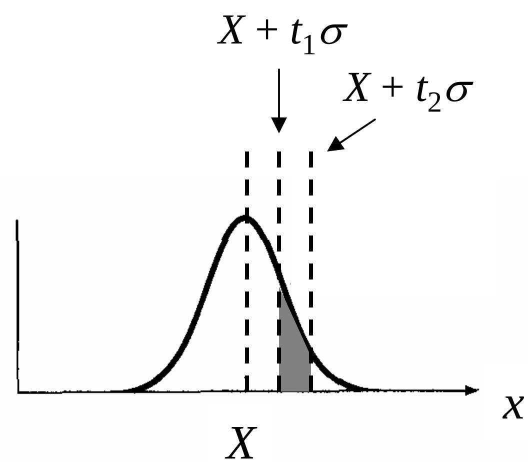
$$t = \frac{x - X}{\sigma}$$

Table B

$$\text{Prob}(X \le x \le X + t\sigma) = \int_X^{X+t\sigma} G_{X,\sigma}(x)dx = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-z^2/2} dz$$

The probability of a measurement to be between $X + t_1\sigma$ and $X + t_2\sigma$

$$\text{Prob}(X + t_1\sigma \le x \le X + t_2\sigma) = \frac{1}{\sqrt{2\pi}} \int_0^{t_2} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{t_1} e^{-z^2/2} dz$$



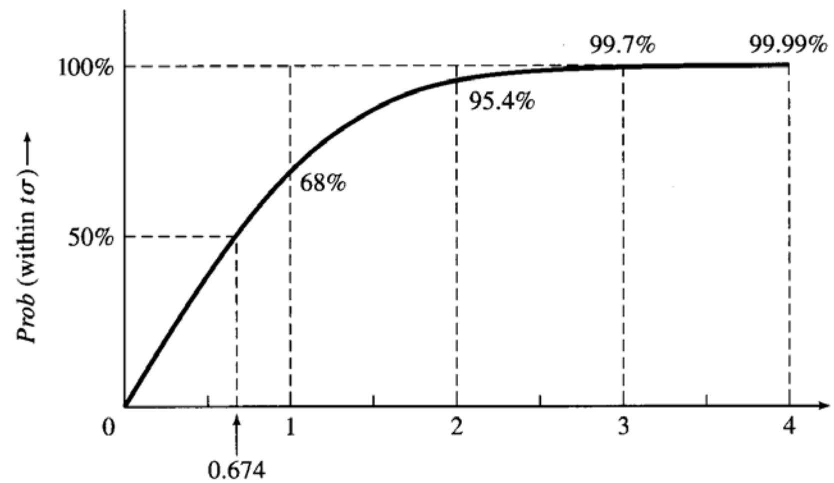
$$t_1 = \frac{x_1 - X}{\sigma} = \frac{11 - 10}{1} = 1$$

$$t_2 = \frac{x_2 - X}{\sigma} = \frac{12 - 10}{1} = 2$$

$$\text{Prob}(X + \sigma \le x \le X + 2\sigma) =$$

$$\frac{1}{\sqrt{2\pi}} \int_0^{2} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{1} e^{-z^2/2} dz =$$

$48\% - 34\% = \underline{14\%}$



Acceptability of a Measured Answer

(value of x) = $x_{best} \pm \sigma$

x_{exp} - expected value of x, e.g. based on some theory

x_{best} differs from x_{exp} by t standard deviations → $t = \dfrac{|x_{best} - x_{exp}|}{\sigma}$

$$\text{Prob(outside } t\sigma) = 1 - \text{Prob(within } t\sigma)$$

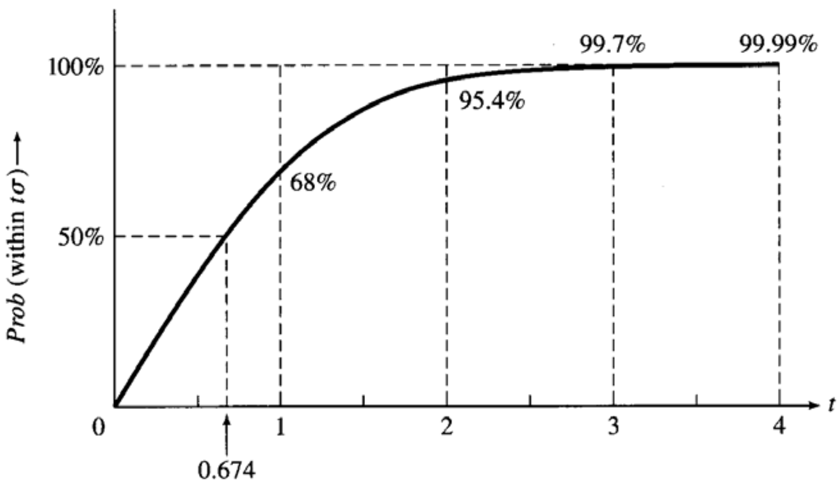
↓

< 5 % - significant discrepancy, $t > 1.96$

< 1 % - highly significant discrepancy, $t > 2.58$

↑

boundary for unreasonable improbability

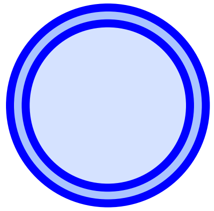


erf(t) – error function

the result is beyond the boundary for unreasonable improbability → the result is unacceptable

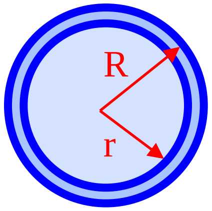
Experiment 2

- Devise a simple, fast, and non-destructive method to measure the variation in thickness of the shell of a large number of racquet balls to determine if the variation in thickness is much less than 10%.



- Devise a method to measure the density of the outer cylinder without damaging the rod so that rods outside 5% tolerance will not be used in a machine.

Moments of Inertia



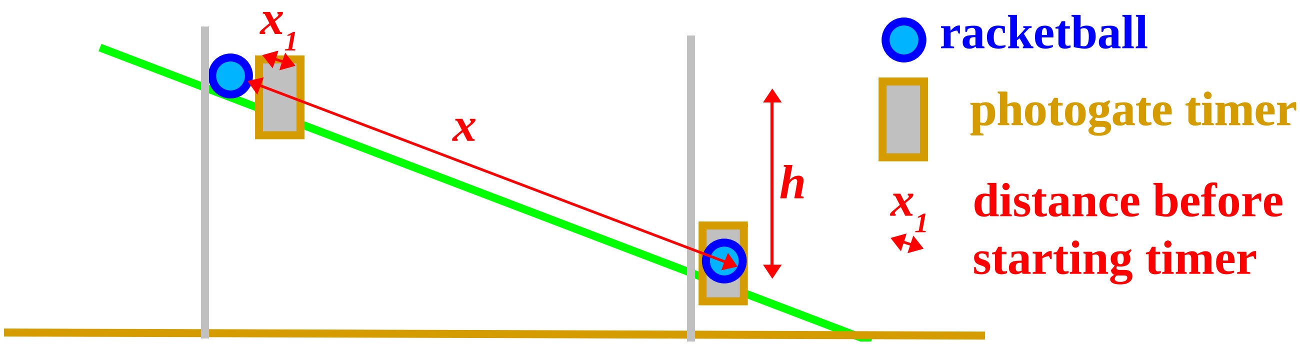
$$I = \frac{2}{5} M \frac{R^5 - r^5}{R^3 - r^3}$$

$$I = \frac{1}{2} m r^2 + \frac{1}{2} M \left(R^2 + r^2 \right)$$

- Both problems can be solved by measuring the mass and moment of inertia of the objects.

- For the balls, we need to measure the variation in thickness.
- For the rods, we need absolute measurements of the density.

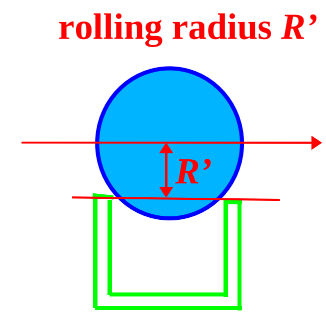
Measuring I by Rolling Objects



1. Measure M and R
2. Using photo gate timer measure the time, t, to travel distance x

$$Mgh = \frac{1}{2}Mv^2 + \frac{1}{2}I\omega^2$$ energy conservation

$$v = R'\omega$$ rolling radius



$$v = \frac{2x}{t}$$ for uniform acceleration

$$Mgh = \frac{1}{2}v^2\left(M + \frac{I}{R'^2}\right)$$

$$gh = \frac{2x^2}{t^2}\left(1 + \frac{I}{MR'^2}\right)$$

$$\frac{I}{MR'^2} = \left(\frac{ght^2}{2x^2} - 1\right)$$

$$\tilde{I} \equiv \frac{I}{MR^2} = \frac{R'^2}{R^2}\left(\frac{ght^2}{2x^2} - 1\right)$$

Measuring the Variation in Thickness of the Shell

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum(x_i - \overline{x})^2}$$

statistical analysis

- 1. Measure rolling time of one ball many times to determine the measurement error in *t* → $\sigma_{\text{measruement}}$
- 2. Measure rolling time of many balls to determine the total spread in *t* → σ_{total}
- 3. Calculate the spread in time due to ball manufacture, $\sigma_{\text{manufacture}}$, by subtracting the measurement error → $\sigma_{\text{total}} = \sigma_{\text{manufacture}} \oplus \sigma_{\text{measruement}}$
- 4. Propagate error on *t* into error on *I* and then into error on thickness *d*

$\sigma_t \longrightarrow \sigma_I \longrightarrow \sigma_d$

variation in *t* → variation in *I* → variation in *d*

Propagate Error from *I* to *d*

$$I = \frac{2}{5} M \frac{R^5 - r^5}{R^3 - r^3}$$

$$z \equiv \frac{r}{R} \approx \frac{28.25 - 4.5 \text{ mm}}{28.25 \text{ mm}} \approx 0.841$$

measured thickness and radius for one ball

d=4.5 mm   *R*=28.25 mm

d=*R*-*r*

$$\tilde{I}(0.841) \equiv \frac{I}{MR^2} = \frac{2}{5}\frac{1-z^5}{1-z^3} \approx 0.571892$$

$$\tilde{I}(0.840) \equiv \frac{I}{MR^2} = \frac{2}{5}\frac{1-z^5}{1-z^3} \approx 0.571366$$

$$\frac{\partial z}{\partial \tilde{I}} = \frac{0.841 - 0.840}{0.571892 - 0.571366} = \frac{0.001}{0.00526} = 1.901$$

← $\delta z \leftrightarrow \delta I$ numerically

$$\frac{\sigma_d}{d} = \frac{\sigma_r}{d} = \frac{R\sigma_z}{d} = \frac{R\tilde{I}}{d}\frac{\partial z}{\partial \tilde{I}}\frac{\sigma_{\tilde{I}}}{\tilde{I}} \approx \frac{(28.25 \text{ mm})(0.572)}{4.5 \text{ mm}}(1.901)\frac{\sigma_{\tilde{I}}}{\tilde{I}} = 6.826\frac{\sigma_{\tilde{I}}}{\tilde{I}} \approx 6.8\frac{\sigma_{\tilde{I}}}{\tilde{I}}$$

$$\frac{\sigma_d}{d} \approx 6.8\frac{\sigma_{\tilde{I}}}{\tilde{I}}$$

Propagate Error from *t* to *I*

from previous page $\tilde{I} = \frac{I}{MR^2} = \frac{R'^2}{R^2}\left(\frac{ght^2}{2x^2} - 1\right) \approx 0.572$

propagate error $\sigma_{\tilde{I}} = \frac{\partial \tilde{I}}{\partial t}\sigma_t = \frac{R'^2}{R^2}\left(\frac{ght}{x^2}\right)\sigma_t = \frac{2}{t}\left(\tilde{I} + \frac{R'^2}{R^2}\right)\sigma_t$

work out fractional error numerically (plug $\tilde{I} \approx 0.572$)

$$\frac{\sigma_{\tilde{I}}}{\tilde{I}} \approx \frac{\frac{2}{t}\left(0.572 + \frac{R'^2}{R^2}\right)}{0.572}\sigma_t \approx 4\frac{\sigma_t}{t}$$

$$\frac{\sigma_{\tilde{I}}}{\tilde{I}} \approx 4\frac{\sigma_t}{t}$$

$$\frac{\sigma_d}{d} \approx 6.8\frac{\sigma_{\tilde{I}}}{\tilde{I}} \approx 27\frac{\sigma_t}{t}$$