# Random Variables

Numeric outcomes of a random process.

outcomes $\xrightarrow{R.V.}$ Numeric Values

Probability Distribution, mean, standard deviation...
Theoretical Results

Most R.V.'s can be associated w/ a probability dist. The ones we will work with will have a mean and std dev.

The value of a random variable will vary from trial to trial.
Each trial or sampled value is a random variable.

We take sample results and combine to find [totals, means, variances, and standard deviations.]

summary statistics

**Section 3.4 - Notation**

**So far we can define random variables: discrete and continuous**

| | X | Y |
|---|---|---|
| probability | $P(X=x)$<br>$P(X \le x)$ | same idea. |
| mean | $\mu_X$ | ⋮ |
| variance | $\sigma_X^2$ | ⋮ |
| standard deviation | $\sigma_X$ | ⋮ |

We sample from popluations.

Suppose a machine produces metal rods. The random variable X is the length.

X = length of the metal rod.

Continuous R.V.!

$\mu_X = 10$ cm $\qquad \bar{x} = 7$ cm

$\sigma_X = 1$ cm

**Consider Random Variable X:**

**Add a constant to create a new random variable.**
**What happens to the mean and standard deviation?**

**Let X represent a randomly selected annual salary at a particular company.**

**Suppose each employee at the company gets a $5000 end-of-year bonus.**
**This is a new random variable!**

X = annual salary

Continuous? If you know everyone's salary.

$Y = X + \$5000$

Y = annual salary w/ bonus

$\mu_x$ = average salary.

$\mu_y = \mu_x + \$5000$ shifted up!

$\sigma_y = \sigma_x$ No change in spread.

**Consider Random Variable X:**
**Multiply by a constant to create a new random variable.**
**What happens to the mean and standard deviation?**

X is some R.V.

$Y = cX$

$\mu_x = \Sigma x \cdot P(X) \rightarrow \mu_y = \Sigma y \cdot P(Y)$

$\mu_y = \Sigma cx \cdot P(X) = c \Sigma x \cdot P(X) = c\mu_x$

$\sigma_x^2 = \Sigma x^2 P(x) - \mu_x^2$

$\sigma_y^2 = \Sigma (cx)^2 P(x) - (c\mu_x)^2$

$= c^2 \Sigma x^2 P(x) - c^2 \mu_x^2$

$= c^2 \sigma_x^2 \qquad \sigma_y = |c| \sigma_x$

**Let X represent the diameter of a randomly selected bolt.**

**If we are interested in the circumference then we can create a new random variable and predict its mean, variance, and standard deviation.**

X = Diameter of a randomly selected bolt

$Y = \pi X$

$\mu_y = \pi \mu_x \qquad \sigma_y^2 = \pi^2 \sigma_x^2$

$\sigma_y = \pi \sigma_x$

**Linear Combinations of INDEPENDENT Random Variables:**

How is standard deviation affected?

$X_1, X_2, ..., X_n$

$\sigma_1^2 \quad \sigma_2^2 \quad \sigma_n^2$

Sum= $X_1+X_2+ ...+X_n$

Variance of the* sum = sum of the variances ♡

$$\sigma_{sum}^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2$$

$$\sigma_{sum} = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}$$

$X_1, X_2, ..., X_n$

$y = cX_1+cX_2+ ...+cX_n$

$$\sigma_y^2 = c^2\sigma_1^2 + c^2\sigma_2^2 + ... + c^2\sigma_n^2$$

$$= c^2 \cdot \sigma_{sum}^2$$

$$\sigma_y = |c|\sigma_{sum}$$

X and Y

$Z = X - Y = X + (-1)Y$

$$\sigma_z^2 = \sigma_x^2 + (-1)^2\sigma_y^2$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2$$

$$\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$$

Wood beams are made up of three layers. Layers are a simple random sample from a population of boards.

$\mu_x = 1.2$ in

$\sigma_x = 0.1$ in

$y = X_1 + X_2 + X_3$

$\mu_y = \mu_x + \mu_x + \mu_x = 3 \cdot \mu_x = 3(1.2 \text{ in})$

$= 3.6$ inches

$$\sigma_y = \sqrt{\sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_3}^2}$$

$$= \sqrt{3(0.1)^2}$$

$\approx 0.17$ inches.

Example: Three steel rods are going to be assembled end to end.

Length of the first component: $\mu_x = 5.1$cm $\sigma_x = 0.2$cm

Length of the second component: $\mu_y = 4.3$ cm $\sigma_y = 0.1$cm

Length of the third component: $\mu_z = 0.5$ cm $\sigma_z = 0.09$cm

Once assembled, find an estimate of the mean total length and the uncertainty (standard deviation) of the total length.

$$T = X + Y + Z$$

$$\mu_T = \mu_x + \mu_y + \mu_z$$

$$= 9.9 \text{ cm}$$

$$\sigma_T = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \approx 0.241 \text{ cm}$$

Often in statistics we take a sample (SRS-independent) and we find the mean of the values in the sample.

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$= \frac{1}{n}X_1 + \frac{1}{n}X_2 + \ldots + \frac{1}{n}X_n$$

$$\mu_{\bar{x}} = \frac{1}{n}\mu_x + \frac{1}{n}\mu_x + \ldots + \frac{1}{n}\mu_x$$

$$= n \cdot \frac{1}{n}\mu_x$$

$$= \mu_x$$

$$\sigma_{\bar{x}}^2 = \frac{1}{n^2}\sigma_x^2 + \frac{1}{n^2}\sigma_x^2 + \ldots + \frac{1}{n^2}\sigma_x^2$$

$$= n \cdot \frac{1}{n^2}\sigma_x^2$$

$$= \frac{\sigma_x^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

So to sum up...this is a big deal!!

**How can we reduce standard deviation (uncertainty)...?**

**Take several INDEPENDENT measurements and find the mean.**

**We create a new random variable that has its own mean and standard deviation.**

New R.V. $\rightarrow \bar{x}$ :)

$\mu_{\bar{x}} = \mu_x \qquad \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

**The mean of several measurements has the same accuracy as a single measurement BUT is more precise.**

**Example: I know a scale produces measurements with uncertainty 0.2g.**

**We want to use the scale to measure the mass of an object.**

**(1) If we make ONE measurement, then the uncertainty does not change.**

$n = 1 \qquad \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{0.2g}{\sqrt{1}} = 0.2g$

**(2) If we make five measurements and find the sample mean, then we can decrease the uncertainty.**

$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{5}} = \frac{0.2g}{\sqrt{5}} = 0.089g$

3.4 HW is due Monday