Applications of the SDM Theorem.

Let X be a population with mean μ and std. dev σ. Fix a sample size $n \geq 30$. Then (under some mild conditions) the rv

$$\overline{X} = \{\text{all means of size-}n\text{ samples from } X\}$$

is roughly $N(\bar{\mu}, \bar{\sigma})$, where

$$\bar{\mu} = \mu \quad \text{and} \quad \bar{\sigma} = \frac{\sigma}{\sqrt{n}}.$$

So: larger n means smaller $\bar{\sigma}$, which means the sample means are more tightly grouped around the mean $\bar{\mu} = \mu$, which means a given sample mean $\overline{x}$ is more likely to closely reflect μ.

So: more data is quantifiably better.

SDM application #1: hypothesis testing.

Example.

A random sample of $n = 130$ people are tested for body temperature. For this sample, the mean temperature is computed to be $\overline{x} = 98.246$, with sample std dev $s = 0.733$.

Test, "at the 98% level," the hypothesis that mean body temperature is 98.6.

Solution.

Let μ denote the actual mean body temperature of all people.

Write μ_0 for the commonly accepted mean body temperature $\mu_0 = 98.6$.

We test the "null hypothesis"

$$H_0 : \mu = \mu_0$$

against the "alternative hypothesis"

$$H_A : \mu \neq \mu_0,$$

at the 98% level. Here's how.

(1) If μ really does equal μ_0 then, by SDM theorem and NISNID fact,

$$\frac{\overline{X} - \overline{\mu_0}}{\overline{\sigma}} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

is roughly $N(0,1)$. Now we don't know σ, so we approximate it by s. So:

(2) If $\mu = \mu_0$, then the rv

$$Z = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} \quad \text{is roughly } N(0,1).$$

(3) From facts about $N(0,1)$, then, we know that

$$P(-2.33 < Z < 2.33) = 98\%.$$

That is: assuming H_0, the "test statistic"

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

will lie inside the interval $(-2.33, 2.33)$ 98% of the time (that is, for 98% of all size-n samples from X). So z will be <u>outside</u> $(-2.33, 2.33)$ only 2% of the time.

(4) So if we compute z, and it <u>is</u> outside this interval – that is, if $|z| \geq 2.33$ – then we suspect that maybe μ is <u>not</u> equal to μ_0. In this case, we <u>reject</u> $H_0: \mu = \mu_0$, and <u>accept</u> $H_A: \mu \neq \mu_0$, at the 98% level (that is, with "at least 98% confidence").

(5) In the present case, we compute:

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{98.246 - 98.6}{0.773/\sqrt{130}} = -5.22.$$

Since $|-5.22| \geq 2.33$, we <u>reject</u> $H_0: \mu = 98.6$, and accept $H_A: \mu \neq 98.6$, at the 98% level.

<u>Notes.</u>

A) "Do not reject H_0" is not the same as "accept H_0." It just means there's insufficient evidence to accept it.

B) Also common are tests at the 95% or 99% level. Here, we compare $|z|$ to 1.96 or 2.58 respectively.

(c) The idea is: the more confident you want or need to be in H_A, the further $\bar{x}$ should be from μ_0, so the larger $|z|$ should be.