Inference for a single mean


My Goodness
My
Guinness

Meet William Sealy Gosset.

Problem: A batch of beer should have a fixed [chemical level related to barley] in order to be of good quality. Can you test a small number of barrels and infer if the entire batch is of good enough quality?

VOLUME VI MARCH, 1908 No. 1

BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

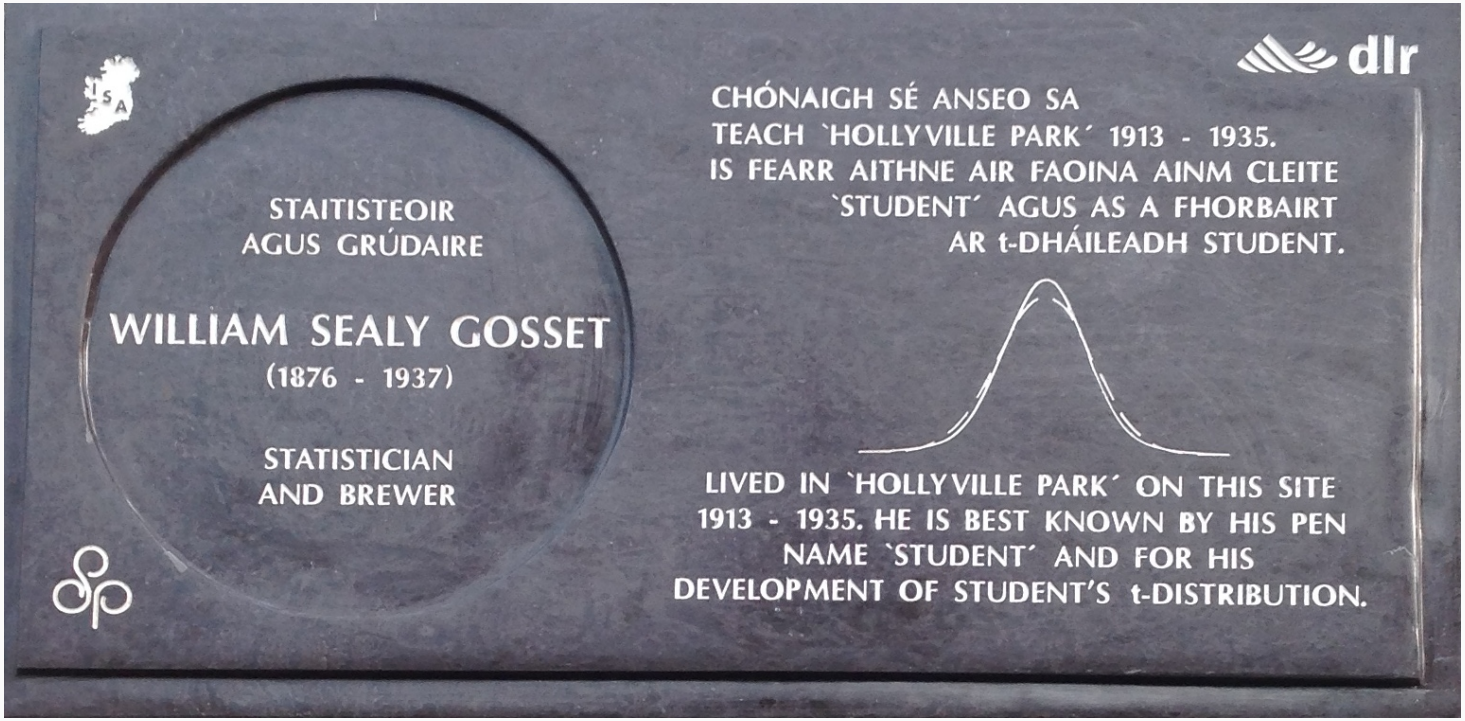
BY STUDENT.

Introduction.

ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

Now any series of experiments is only of value in so far as it enables us to form a judgment as to the statistical constants of the population to which the experiments belong. In a great number of cases the question finally turns on the value of a mean, either directly, or as the mean difference between the two quantities.

If the number of experiments be very large, we may have precise information


dlr
STAITISTEOIR
AGUS GRÚDAIRE
WILLIAM SEALY GOSSET
(1876 - 1937)
STATISTICIAN
AND BREWER
CHÓNAIGH SÉ ANSEO SA
TEACH `HOLLYVILLE PARK´ 1913 - 1935.
IS FEARR AITHNE AIR FAOINA AINM CLEITE
`STUDENT´ AGUS AS A FHORBAIRT
AR t-DHÁILEADH STUDENT.
LIVED IN `HOLLYVILLE PARK´ ON THIS SITE
1913 - 1935. HE IS BEST KNOWN BY HIS PEN
NAME `STUDENT´ AND FOR HIS
DEVELOPMENT OF STUDENT'S t-DISTRIBUTION.

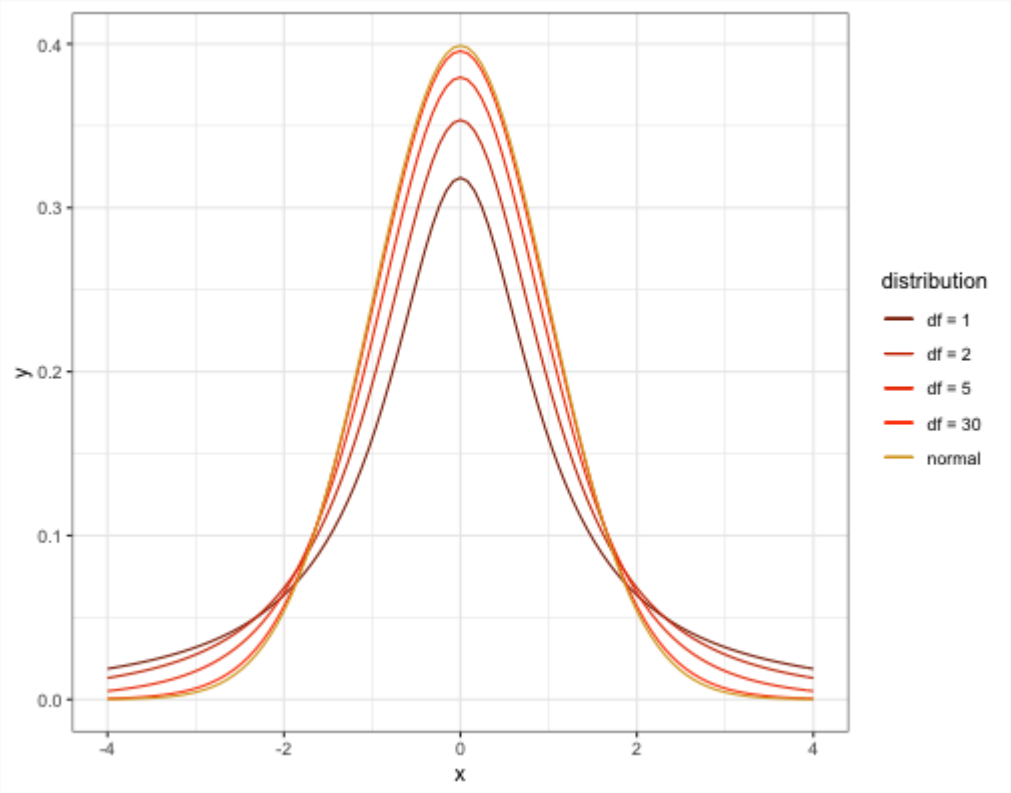
The t distribution

Used to estimate the mean when you have a small sample drawn from a nearly normal population.

Conditions

- Independent observations ($n < 0.1N$)
- Nearly normal population distribution
  - Check distribution of the sample as a proxy

t versus normal



The t has heavier tails than the normal distribution.

Degrees of Freedom

The number of parameters that are free to vary, without violating any constraint imposed on it.

Parameters

μ

Since $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, one of our *observations* is constrained, leaving $n - 1$ that are free to vary.

$$df = n - 1$$

Hypothesis testing

1. State hypotheses: e.g. $H_0 : \mu = \mu_0$ versus $H_A : \mu \neq \mu_0$

2. Check conditions

    - Independent observations
    - Nearly normal population

3. Compute observed t-statistic

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

4. Draw picture to assess where t_{obs} falls in $t_{df=n-1}$
5. Compute a (two-tailed) p-value
6. State conclusion

Confidence interval for μ

point estimate $\pm$ margin of error

$$\bar{x} \pm (t^*_{df} \times SE)$$

- $\bar{x}$: point estimate of μ.
- t^*_{df}: critical value that leaves α in the tails of a t with $df = n - 1$.
- SE: standard error of $\bar{x}$, $s/\sqrt{n}$.

Finding p-values and t^*_{df}

```
pt(-2.2, df = 18)
```

```
## [1] 0.0206
```

```
qt(.025, df = 18)
```

```
## [1] -2.1
```

- Use the applet

Example: The Kilogram

How do we know how much a kilogram weighs?

Example: The Kilogram, cont.

Meet the IPK: International Prototype Kilogram, Paris, France.

Example: The Kilogram

Question

The US has two copies of the IPK. Say they make a third copy for Reed. We take 6 measures of our new RPK and get the following:

```
## [1] 0.997 0.976 1.032 1.101 1.004
```

Is it a safe assumption that the RPK weighs the same as the IPK?

RIP IPK

The new kilogram just debuted. It's a massive achievement.

The new definition represents a victory of humankind over chaos in the universe. Really.

By Brian Resnick | @B_resnick | brian@vox.com | Updated May 20, 2019, 8:45am EDT

SHARE

Until now, the definition of a kilogram has been a hunk of metal in Paris

For more than a century, the kilogram had a very simple definition: It was the mass of a hunk of platinum-iridium alloy that's been housed at the International Bureau of Weights and Measures in Sèvres, France **since 1889**.

It's called the International Prototype Kilogram (a.k.a. Big K, or Le Grand K), and it has many copies around the world — including **several at NIST in Gaithersburg, Maryland** — that are used to calibrate scales and make sure the whole world is on one system of measurement.