Tues 2pm

Applied Statistics

Lectures 7

David Barton & Sabine Hauert

Department of Engineering Mathematics

Outline

- Sample mean with unknown variance
- t-distribution

OpenIntro Statistics

Chapter 5, particularly §5.1, §5.2, and §5.3

Back to the sample mean. . .

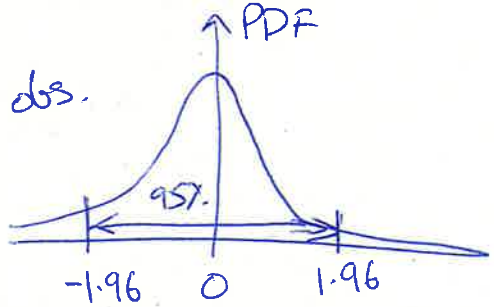
The test statistic for the sample mean makes use of the *population variance* — what happens when the variance is unknown? We now have (almost) all the ingredients.

sample mean
population mean

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

number of obs.

population variance



becomes

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim ???$$

We now have a normally distributed number divided by a χ^2 distributed number!

sample standard deviation (variance)

What is the resulting distribution?

Time for a pint Guinness

Student's t-distribution

Student's t-distribution takes its name from William Sealy Gosset writing under the pseudonym Student. He worked for the Guiness Brewery in Dublin and was interested in the properties of small samples (e.g., barley).

Definition (Test statistic for the sample mean (normal))

The test statistic for the sample mean of a normally distributed random sample with *estimated variance* is

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim T_{n-1}.$$

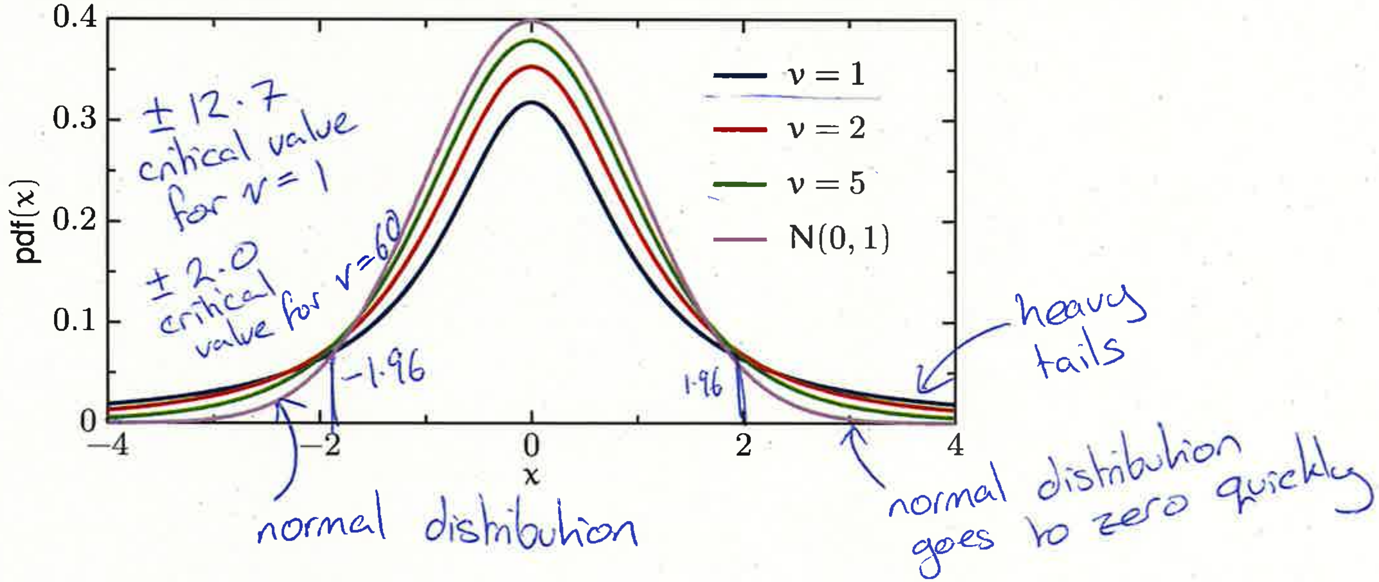
As $n \to \infty$ we have that $T_{n-1} \rightsquigarrow N(0, 1)$. (How large is large?)

Student's t-distribution

The PDF of Student's t-distribution is

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + x^{\frac{2}{\nu}}\right)^{\left(-\frac{\nu+1}{2}\right)}$$

ν is the number of degrees of freedom and Γ is the gamma function.



When to use the t-distribution

Students are often confused about when to use the t-distribution; the choice is relatively simple —

given in the hypothesis ?

If you *know or are given* the population variance, use the normal distribution (i.e., Z-score)

If you don't know the population variance and have to *calculate the sample variance*, use a t-test.

- For large sample sizes ($n \geqslant 30$) the t-distribution and normal distribution are very similar so you can use either
- For small sample sizes ($n < 30$) you must use the t-distribution.

$s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$

Exercise

A retailer states that the average load of goods transported by a company is 25 tonnes. The transport company measures 5 such loads and finds the values

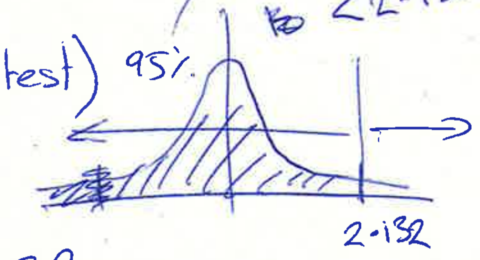
25.8   36.6   26.3   21.8   27.2

They believe the mean weight $\boxed{\text{exceeds}}$ 25. By evaluating a T-test statistic, establish whether this distribution is consistent with the hypothesis that the mean weight is actually 25, as stated, using a significance level of 0.05 (i.e. 5%).

$H_0: \mu = 25 \qquad H_1: \mu > 25$

$T = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim T_{n-1}(0.05)$ (one-tailed test)

95% of all model outputs ~~to~~ < 2.132



$\bar{x} = 27.54 \qquad s = 5.47$

$T = \frac{27.54 - 25}{5.47/\sqrt{5}} = 1.038 < 2.132$

$\therefore$ cannot reject H_0

When is something insignificant significant?

Definition (Statistical significance)

Statistical significance occurs when the p-value is lower than the *a priori* chosen significance level (often 5%). Equivalently, it is when the measured results lie in the rejection region for a hypothesis test.

Statistical significance tends to occur when either there are (a) many samples and a small deviation from the hypothesis or (b) a few samples and a large deviation from the hypothesis.

Don't confuse statistical significance for practical significance!

- Is a small deviation from the hypothesis of practical significance?
- A moderate deviation not be detected with only a few samples.

Failing to reject the null hypothesis

The 1973 oil crisis led to traffic agencies considering whether to allow vehicles to turn right at a red light (USA). There were significant safety concerns and a number of studies performed.

In one a consultant in Virginia conducted a before-and-after study of 20 intersections.

Before there were 308 accidents and
Afterwards there were 337 accidents in a similar time frame.

This result *was not statistically significant*. The final report to the governor stated

> *we can discern no significant hazard to motorists or pedestrians from the implementation*

Notice how practical significant has been equated to statistical significance.

Failing to reject the null hypothesis

These studies were *underpowered* — they were not sufficiently sensitive to the phenomenon they were trying to measure!

Subsequent studies have found

- 20% increase in collisions,
- 60% more pedestrians being run over, and
- twice as many bicyclists being struck!

[Source: Statistics done wrong, Reinhart]

Just because you can't reject your hypothesis doesn't mean it's right!

Quote of the day

Elizur Wright

While nothing is more uncertain than a single life, nothing is more certain than the average duration of a thousand lives.

Exercises

- A selection from §5.6 in OpenIntro Statistics