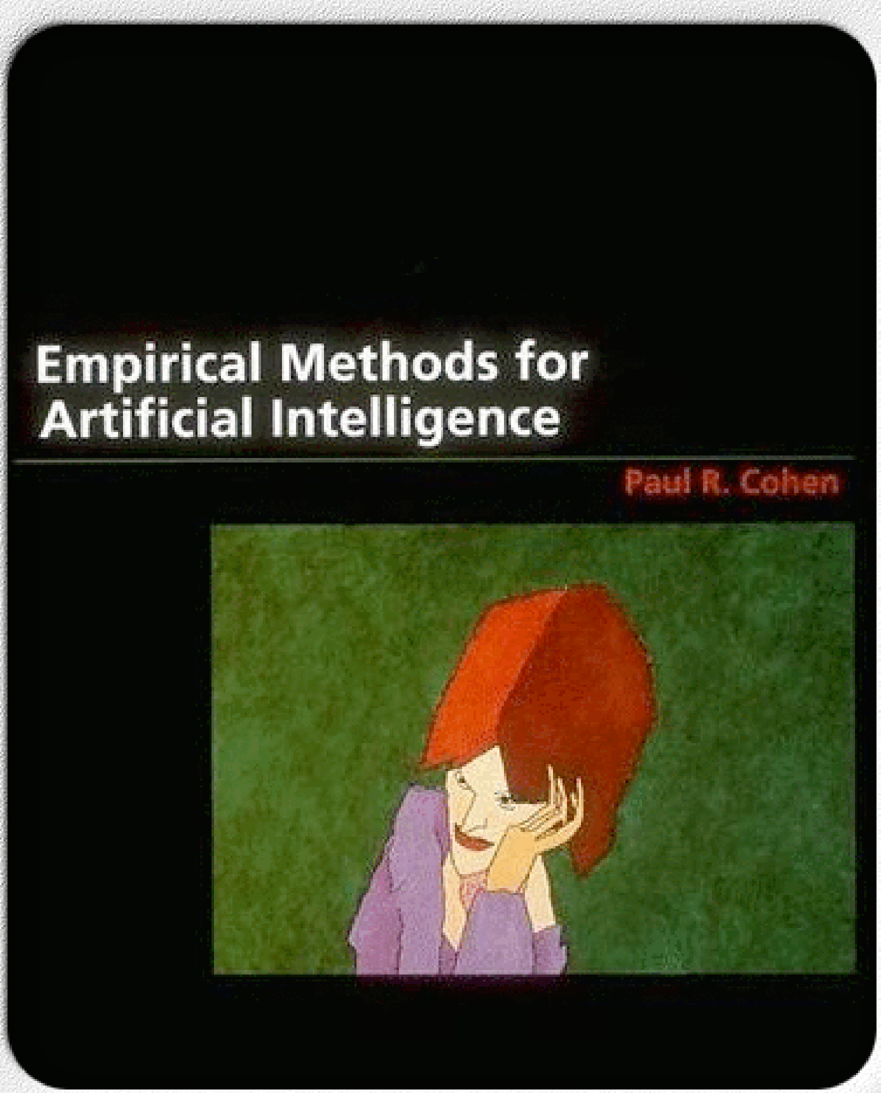
Optimization weekly meeting
Apr 27, 2006



Chapter 4:
Hypothesis Testing and Estimation

IRIDIA - ULB
Max MANFRIN

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations

Statistical inference is inference from statistics (functions on samples like $\bar{x}, s^2, s$) to parameters (functions on populations like μ_x, σ^2, σ).

Need to focus on relationship between the number of data in a sample, the variance of the data, and our confidence in conclusions.

Hypothesis Testing

Answer a yes-or-no question about a population and assess the probability that the answer is wrong.

Parameter Estimation

Estimate the true value of a parameter given a statistic.

Sec 4.1

Statistical inference is inference from statistics (functions on samples like $\bar{x}, s^2, s$) to parameters (functions on populations like μ_x, σ^2, σ).

Need to focus on relationship between the number of data in a sample, the variance of the data, and our confidence in conclusions.

Hypothesis Testing

How likely is a sample result given an assumption about the population?

Hy

An ... estion about a population and assess the ... answer is wrong.

Parameter Estimation

Estimate the true value of a parameter given a statistic.

Sec 4.1

Statistical inference is inference from statistics (functions on samples like $\bar{x}, s^2, s$) to parameters (functions on populations like μ_x, σ^2, σ).

Need to focus on relationship between the number of data in a sample, the variance of the data, and our confidence in conclusions.

Hypothesis Testing

How likely is a sample result given an assumption about the population?

Parameter Estimation

What is the most likely value of a parameter, and what are likely bounds on the value of a parameter given a statistic?

Hy

An … estion about … ssess the … answer is wro

Parameter Estimation

Estimate the true value of a paramete … statistic.

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations

The question of interest is simplified into two competing claims / hypothesis (the hypothesis are often statements about population parameters): the **null hypothesis**, denoted **H_0** and the **alternative hypothesis**, denoted **H_1**.

Special consideration is given to H_0 because it relates to the statement being tested, whereas H_1 relates to the statement to be accepted if/when H_0 is rejected.

Statistical hypothesis testing doesn't prove H_0 true or false; it bounds the probability of incorrectly asserting - based on a sample result - that H_0 is false.

The difficult part is establishing the probability of incorrectly rejecting H_0.

Hypothesis Testing

How likely is a sample result given an assumption about the population?

Sec 4.2

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations

A sampling distribution is the distribution of a statistic calculated from all possible samples of a given size, drawn from a given population.

Most sampling distribution are estimated or determined analytically, not constructed, because the population from which samples are drawn are very large or infinite.

If one draws many samples of a given size from a population and calculate a statistic, then the resulting empirical sampling distribution of the statistic will probably be quite similar to the theoretical sampling distribution.

Hypothesis Testing

How likely is a sample result given an assumption about the population?

Sec 4.3.1

Classical statistical methods calculate sampling distribution exactly or estimate them analytically.

Exact Sampling Distribution:
The sampling distribution of the proportion can be calculated exactly (e.g. tossing a coin. Is the coin fair?)

All possible proportions are: $\frac{0}{N}, \frac{1}{N}, \ldots, \frac{N}{N}$

The probability distribution over these values is the binomial distribution. The probability of a particular sample proportion is:

$$\frac{N!}{i!(N-i)!} r^i (1-r)^{N-i}$$

where N is the number of tosses, r is the probability that a single toss will land head, and i is the number of heads observed.

Sec 4.3.2

Classical statistical methods calculate sampling distribution exactly or estimate them analytically.

Exact Sampling Distribution:

The sampling dis... ...portion can be calculated exactly (e.g. tossing

All possible propo... $\ldots, \frac{N}{N}$

The probability d... ...se values is the binomial distribution. Th... a particular sample proportion is:

$N-i$

Continuous probability distributions

While discrete probability distributions provide probabilities for exact outcomes, continuous prob. distr. provide probabilities for ranges of outcomes.

In cont. distr. the prob. of a particular outcome is zero.

where N is the number of tosses, r is the probability that a single toss will land head, and i is the number of heads observed.

Sec 4.3.2

Estimated Sampling Distribution:
The sampling distribution of the mean can be estimated

Central Limit Theorem
The sampling distribution of the mean of samples of size N approaches a normal distribution as N increase. If the samples are drawn from a population with mean μ and a standard deviation σ, then the mean of the sampling distribution is μ and its standard deviation is $\sigma/\sqrt{N}$.
These statements hold irrespective of the shape of the population distribution from which the samples are drawn.

If you draw samples from any population you like, and provided the samples are large (usually means $N \geq 30$), the sampling distribution of the sample mean is normal.

Sampling distribution demo:
http://www.ruf.rice.edu/~lane/stat_sim/sampling_dist/index.html

Sec 4.3.2

The Standard Error of the Mean : the standard deviation of the sampling distribution of the mean; is denoted $\sigma_{\bar{x}}$

The standard deviation of the sampling distribution represents uncertainty about μ. Formally

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

where σ is the population standard deviation.

According to the central limit theorem, the mean of the sampling distribution approaches the population mean μ as N increases.

Standard errors of almost all statistics are hard to calculate or have no closed form expression.

Sec 4.3.2

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations

Z Test: take a sample, and transform it into Z score

$$z_i = \frac{x_i - \bar{x}}{s}$$

Z score transformation does not change the underlying distribution, just the scale.

This does allow for certain useful operations: comparison of relative position in different samples, combination of scores on different scales.

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

Essentialy it asks: **Given a population with a certain mean, how likely is it I could draw a sample with my given sample mean?**

Z score falls in the lower or upper 5% of the sampling distribution, it's considered adequate evidence against H_0.

Z score falls in the lower or upper 2.5% of the sampling distribution, it's considered good evidence against H_0.

Z score falls in the lower or upper 0.5% of the sampling distribution, it's considered strong evidence against H_0.

Sec 4.4

Z Test: take a sample, and transform it into Z score

$$z_i = \frac{x_i - \bar{x}}{s}$$

Z score transformation does not change the underlying distribution, just the scale.

This does allow for ... operations: comparison of ... les, combination of scores

... u

... on with a certain mean, ... e with my given sample ...

Z s... sampling distribution, it's considered ade...

Z sc... lower or upper 2.5% of the sampling distribution, it's considered good evidence against H_0.

Z score falls in the lower or upper 0.5% of the sampling distribution, it's considered strong evidence against H_0.

Z Test

- It estimates the sampling distribution of the mean.
- It transforms the sampling distribution into a standard normal distribution.
- It expresses the sample mean as Z standard deviations distance from its expectation under the null hypothesis.

Sec 4.4

Critical value: the value corresponding to a given significance level. This cutoff value determines the boundary between those samples resulting in a test statistic that leads to rejecting H_0 and those that lead to a decision not to reject H_0.

p value: the probability of obtaining a particular sample result given the null hypothesis.

(By convention, one usually doesn't reject the null hypothesis unless $p < 0.05$. This is the largest bound on p that most researchers will call statistically significant.)

An exact value for p is not the exact probability of the sample result under H_0 unless the sampling distribution is discrete.

Sec 4.4

When σ is unknown: it's rarely the case that we know the population standard deviation, so we have to estimate it from the sample standard deviation: $\hat{\sigma} = s$

Similarly, the standard error is estimated as follows:

$$\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{N}} = \frac{s}{\sqrt{N}}$$

Z Test can be run as:

$$Z = \frac{\bar{x} - \mu}{\hat{\sigma}_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{N}}}$$

Sec 4.4

When all population parameters are unknown: a common tactic is to compare a sample to an imagined null hypothesis distribution.

'Suppose a population does exist and its mean is x; could my sample have come from it?'

In practice we use s, the sampled standard deviation, to estimate σ and $\sigma_{\bar{x}}$ and to run the test as before.

Sec 4.4

If N ≥ 30, we can be sure that the sampling distribution is normal. But what about smaller samples, like N=20, N=10 or even N=5 ?

t distribution: the sampling distribution of the mean for small N; it looks a lot like the normal distribution, but more of the mass of the distribution is in the tails.

A sample results that's highly improbable when matched to a normal distribution is more likely when matched to the t distribution.

t test can be run as:

$$t = \frac{\bar{x} - \mu}{\hat{\sigma}_{\bar{x}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{N}}}$$

the same as Z test, except that you compare a t score to the t distribution. There is a t distribution for each N.

Sec 4.4

If N ≥ 30, we can be sure that the sampling distribution is normal. But what about smaller samples, like N=20, N=10 or even N=5 ?

t distribution: the sampling distribution of the mean for small N; it looks a lot like the normal distribution, but more of the mass of the distribution is in the tails.

A sample results that's highly
normal distribution is more
distribution.

t test can be run as:

$$t = \frac{\bar{x}-\mu}{\hat{\sigma}_{\bar{x}}} =$$

t Test

it depends on one important assumption: The distribution from which the sample is drawn is normal.

Whether the data points are normally distributed can be assessed by a normality test, such as Kolmogorov-Smimov or Shapiro-Wilk.

the same as Z test, except that you compare a t score to the t distribution. There is a t distribution for each N.

Sec 4.4

Three types of t tests are commonly run on means:

One-sample t Test: assess whether a sample was drawn from a population with known mean μ.

One-sample t Test: assess whether a hypothetical μ could be the mean of the population from which a sample was drawn.

Two-sample t Test: assess whether the means of two normally distributed populations are equal.
(we compare the means of two samples to see whether they could have been drawn from populations with equal means).

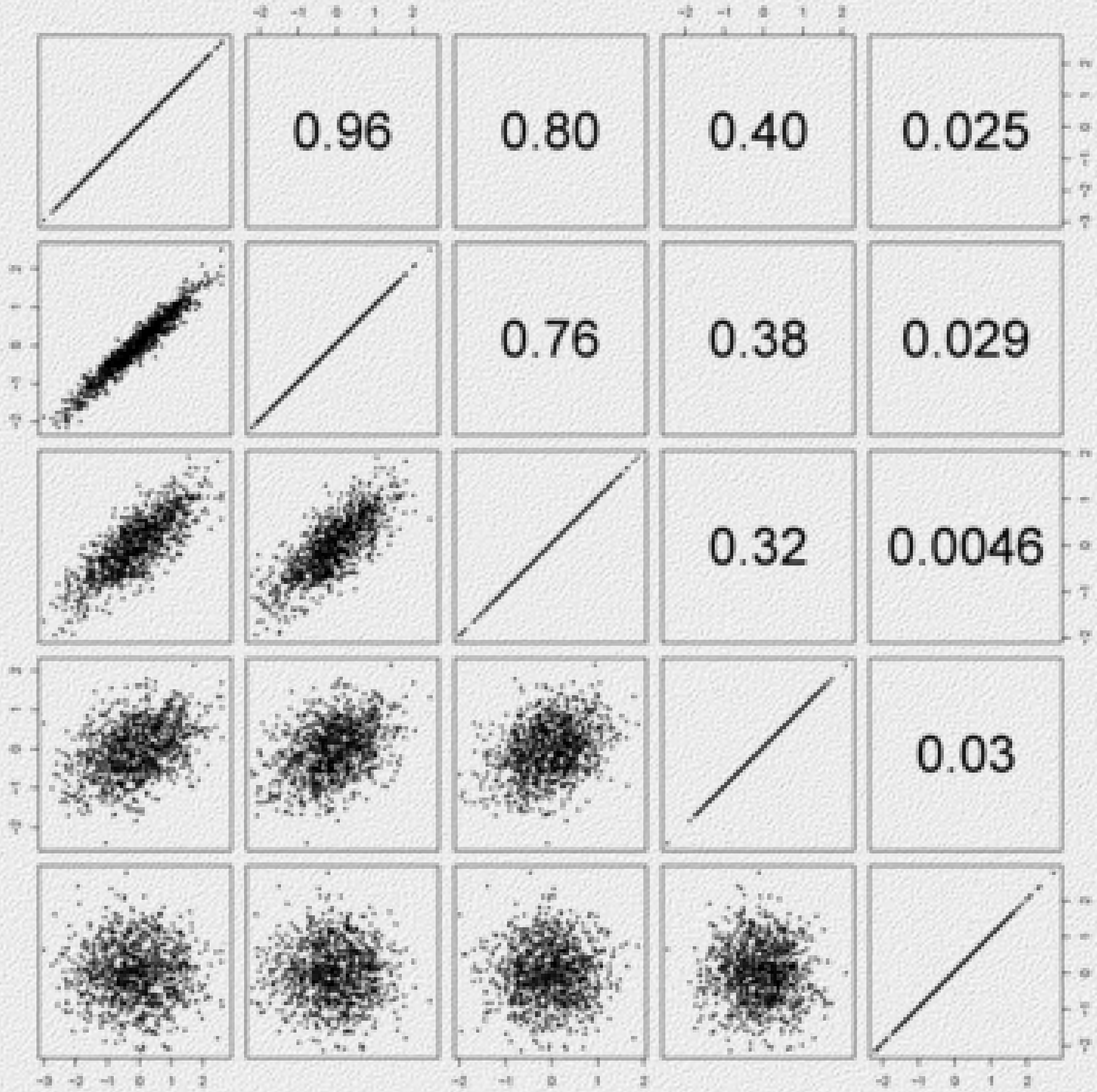
There are different versions of the t Test depending on whether the two samples are independent of each other or paired, so that each member of one sample has a unique relationship with a particular member of the other sample.

Sec 4.4

Outline

part 1

- Statistical Inference
- Introduction to Hypothesis Testing
- Sampling Distributions and Hypothesis Testing Strategy
- Test of Hypotheses about Means
- Hypothesis about Correlations



Linear correlations between 1000 pairs of numbers. The data are graphed on the lower left and their correlation coefficients listed on the upper right. Each square in the upper right corresponds to its mirror-image square in the lower left, the "mirror" being the diagonal of the whole array. Each set of points correlates maximally with itself, as shown on the diagonal (all correlations = +1).

Sec 4.5

We might believe, on the basis of a low correlation, that two variables are independent. To be more certain, we want to test the hypothesis that their correlation is zero.

To do so, we need the sampling distribution of the correlation coefficient, which is complicated.

We can transform the correlation into another form that has a nicer sampling distribution.

Fisher's r to z transform: produces a statistic, the sampling distribution of which is approximately normal under the assumption that the variables are normally distributed.

The statistic is: $z(r) = 0.5 \ \ln \ \frac{1+r}{1-r}$

The mean of the sampling distribution is: $z(\rho) = 0.5 \ \ln \ \frac{1+\rho}{1-\rho}$

The estimated standard error is: $\hat{\sigma}_{z(r)} = \frac{1}{\sqrt{n-3}}$

Sec 4.5

We might believe, on the basi … correlation, that two variables are i … certain, we want to test the hypothe … is zero.

Fisher's r to z transform

After the transformation we can run a Z test of the null hypothesis that the correlation of two variables is some number.

To do so, we nee … n of the correlation coefficient, which …

We can transform … into another form that has a nicer sampling distribution.

Fisher's r to z transform: produces a statistic, the sampling distribution of which is approximately normal under the assumption that the variables are normally distributed.

The statistic is: $z(r) = 0.5 \ \ln \frac{1+r}{1-r}$

The mean of the sampling distribution is: $z(\rho) = 0.5 \ \ln \frac{1+\rho}{1-\rho}$

The estimated standard error is: $\hat{\sigma}_{z(r)} = \frac{1}{\sqrt{n-3}}$

Sec 4.5

End of part 1