

0.1 P values

- The P value or calculated probability is the estimated probability of rejecting the null hypothesis (H0) of a study question when that hypothesis is true.
- The null hypothesis is usually an hypothesis of "no difference" e.g. no difference between blood pressures in group A and group B. Define a null hypothesis for each study question clearly before the start of your study.
- The only situation in which you should use a one sided P value is when a large change in an unexpected direction would have absolutely no relevance to your study. This situation is unusual; if you are in any doubt then use a two sided P value.
- The term significance level (alpha) is used to refer to a pre-chosen probability and the term "P value" is used to indicate a probability that you calculate after a given study.
- The alternative hypothesis (H1) is the opposite of the null hypothesis; in plain language terms this is usually the hypothesis you set out to investigate. For example, question is "is there a significant (not due to chance) difference in blood pressures between groups A and B if we give group A the test drug and group B a sugar pill?" and alternative hypothesis is "there is a difference in blood pressures between groups A and B if we give group A the test drug and group B a sugar pill".
- If your P value is less than the chosen significance level then you reject the null hypothesis i.e. accept that your sample gives reasonable evidence to support the alternative hypothesis.
- It does NOT imply a "meaningful" or "important" difference; that is for you to decide when considering the real-world relevance of your result.
- The choice of significance level at which you reject H0 is arbitrary. Conventionally the 5% (less than 1 in 20 chance of being wrong), 1% and 0.1% ($P \leq 0.05$, 0.01 and 0.001) levels have been used. These numbers can give a false sense of security.

In the ideal world, we would be able to define a "perfectly" random sample, the most appropriate test and one definitive conclusion. We simply cannot. What we can do is try to optimise all stages of our research to minimise sources of uncertainty. When presenting P values some groups find it helpful to use the asterisk rating system as well as quoting the P value:

- $P < 0.05$ *
- $P < 0.01$ **
- $P < 0.001$

Most authors refer to statistically significant as $P \leq 0.05$ and statistically highly significant as $P \leq 0.001$ (less than one in a thousand chance of being wrong).

The asterisk system avoids the woolly term "significant". Please note, however, that many statisticians do not like the asterisk rating system when it is used without showing P values. As a rule of thumb, if you can quote an exact P value then do. You might also want to refer to a quoted exact P value as an asterisk in text narrative or tables of contrasts elsewhere in a report.

0.1.1 p-values

- He or she would have no more basis to doubt the validity of the null hypothesis than if p-value had been 0.482. The conclusion would be that the null hypothesis could not be rejected at the 0.05 level.
- In short, this approach is to specify the significance level in advance and use p-value only to determine whether or not the null hypothesis can be rejected at the stated significance level.
- Many statisticians and researchers find this approach to hypothesis testing not only too rigid, but basically illogical. It is very reasonable to have more confidence that the null hypothesis is false with a p-value of 0.0001 than with a p-value of 0.042?

0.1.2 p-values

- The less likely the obtained results (or more extreme results) under the null hypothesis, the more confident one should be that the null hypothesis is false.
- The null hypothesis should not be rejected once and for all. The possibility that it was falsely rejected is always present, and, all else being equal, the lower the p-value, the lower this possibility.
- According to this view, research reports should not contain the p-value, only whether or not the values were significant (at or below the significance level).
- However it is much more reasonable to just report the p-values. That way each reader can make up his or her mind about just how convinced they are that the null hypothesis is false.

0.1.3 p-value

- The p-value (or p-value or probability value is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true.
- The null hypothesis is rejected if the P-value is very small, such as less than 0.05.

0.1.4 p-values

- The null hypothesis either is or is not rejected at the previously stated significance level. Thus, if an experimenter originally stated that he or she was using the $\alpha = 0.05$ significance level and p-value was subsequently calculated to be 0.042, then the person would reject the null hypothesis at the 0.05 level.
- If p-value had been 0.0001 instead of 0.042 then the null hypothesis would still be rejected at the 0.05 significance level.
- The experimenter would not have any basis to be more confident that the null hypothesis was false with a p-value of 0.0001 than with a p-value of 0.041.
- Similarly, if the p had been 0.051 then the experimenter would fail to reject the null hypothesis

0.1.5 The Hypothesis Testing Procedure

We will use both of the following four step procedures for hypothesis testing. The level of significance must be determined in advance. The first procedures is as follows:

- Formally write out the null and alternative hypotheses (already described).
- Compute the *test statistic* - a standardized value of the numerical outcome of an experiment.
- Compute the p-value for that test statistic.
- Make a decision based on the p-value.

0.1.6 Hypothesis Testing and p-values

- In hypothesis tests, the difference between the observed value and the parameter value specified by H_0 is computed and the probability of obtaining a difference this large or large is calculated.
- The probability of obtaining data as extreme, or more extreme, than the expected value under the null hypothesis is called the *p-value*.
- There is often confusion about the precise meaning of the p-value probability computed in a significance test. It is not the probability of the null hypothesis itself.
- Thus, if the probability value is 0.0175, this does not mean that the probability that the null hypothesis is either true or false is 0.0175.
- It means that the probability of obtaining data as different or more different from the null hypothesis as those obtained in the experiment is 0.0175.

0.2 Using P-Values

p-values (Mentioned Previously, but discussed again)

- The p-value (or P-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true.
- The null hypothesis is rejected if the p-value is very small, such as less than 0.05.
- When performing an inference procedure on a computer, it is much more common to use the p-value as a basis for decision, rather than the critical value

p-values

- The less likely the obtained results (or more extreme results) under the null hypothesis, the more confident one should be that the null hypothesis is false. The null hypothesis should not be rejected once and for all. The possibility that it was falsely rejected is always present, and, all else being equal, the lower the p-value, the lower this possibility.
- According to this view, research reports should not contain the p-value, only whether or not the values were significant (at or below the significance level).

- However it is much more reasonable to just report the p-values. That way each reader can make up his or her mind about just how convinced they are that the null hypothesis is false.

The Critical region The critical region (or rejection region) is the set of all values of the test statistic that causes us to reject the null hypothesis.

Test statistics for testing a claim about a mean, when the population variance is known.

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

0.3 p-values

- In hypothesis tests, the difference between the observed value and the parameter value specified by H_0 is computed and the probability of obtaining a difference this large or larger is calculated.
- The probability of obtaining data as extreme, or more extreme, than the expected value under the null hypothesis is called the **p-value**.
- It is not the probability of the null hypothesis itself.
- Suppose if the probability value is 0.0175, this does not mean that the probability that the null hypothesis is either true (or false) is 0.0175.
- the p-value means that the probability of obtaining data as different or more different from the null hypothesis as those obtained in the experiment is 0.0175.
- If the p-value is less than the specified significance level, adjusted for the number of tails, then we reject the null hypothesis.

$$\text{is p-value} \leq \frac{\alpha}{k}?$$

P-values

- The null hypothesis either is or is not rejected at the previously stated significance level. Thus, if an experimenter originally stated that he or she was using the $\alpha = 0.05$ significance level and p-value was subsequently calculated to be 0.042, then the person would reject the null hypothesis at the 0.05 level.
- If p-value had been 0.0001 instead of 0.042 then the null hypothesis would still be rejected at the 0.05 significance level.
- The experimenter would not have any basis to be more confident that the null hypothesis was false with a p-value of 0.0001 than with a p-value of 0.041.
- Similarly, if the p had been 0.051 then the experimenter would fail to reject the null hypothesis

0.4 P values

The P-value (or p-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true. The null hypothesis is rejected if the P-value is very small, such as less than 0.05.

- The less likely the obtained results (or more extreme results) under the null hypothesis, the more confident one should be that the null hypothesis is false. The null hypothesis should not be rejected once and for all. The possibility that it was falsely rejected is always present, and, all else being equal, the lower the p-value, the lower this possibility.
 - According to this view, research reports should not contain the p-value, only whether or not the values were significant (at or below the significance level).
 - However it is much more reasonable to just report the p-values. That way each reader can make up his or her mind about just how convinced they are that the null hypothesis is false.
-
- The P value or calculated probability is the estimated probability of rejecting the null hypothesis (H_0) of a study question when that hypothesis is true.
 - The null hypothesis is usually an hypothesis of "no difference" e.g. no difference between blood pressures in group A and group B. Define a null hypothesis for each study question clearly before the start of your study.
 - The only situation in which you should use a one sided P value is when a large change in an unexpected direction would have absolutely no relevance to your study. This situation is unusual; if you are in any doubt then use a two sided P value.
 - The term significance level (α) is used to refer to a pre-chosen probability and the term "P value" is used to indicate a probability that you calculate after a given study.
 - The alternative hypothesis (H_1) is the opposite of the null hypothesis; in plain language terms this is usually the hypothesis you set out to investigate. For example, question is "is there a significant (not due to chance) difference in blood pressures between groups A and B if we give group A the test drug and group B a sugar pill?" and alternative hypothesis is "there is a difference in blood pressures between groups A and B if we give group A the test drug and group B a sugar pill".
 - If your P value is less than the chosen significance level then you reject the null hypothesis i.e. accept that your sample gives reasonable evidence to support the alternative hypothesis.
 - It does NOT imply a "meaningful" or "important" difference; that is for you to decide when considering the real-world relevance of your result.
 - The choice of significance level at which you reject H_0 is arbitrary. Conventionally the 5% (less than 1 in 20 chance of being wrong), 1% and 0.1% ($P < 0.05$, 0.01 and 0.001) levels have been used. These numbers can give a false sense of security.

In the ideal world, we would be able to define a "perfectly" random sample, the most appropriate test and one definitive conclusion. We simply cannot. What we can do is try to optimise all stages of our research to minimise sources of uncertainty. When presenting P values some groups find it helpful to use the asterisk rating system as well as quoting the P value:

- $P < 0.05$ *
- $P < 0.01$ **
- $P < 0.001$

Most authors refer to statistically significant as $P < 0.05$ and statistically highly significant as $P < 0.001$ (less than one in a thousand chance).

- The asterisk system avoids the woolly term "significant". Please note, however, that many statisticians do not like the asterisk rating system when it is used without showing P values.
- As a rule of thumb, if you can quote an exact P value then do. You might also want to refer to a quoted exact P value as an asterisk in text narrative or tables of contrasts elsewhere in a report.

0.5 p-values

- A practitioner would have no more basis to doubt the validity of the null hypothesis than if p-value had been 0.482. The conclusion would be that the null hypothesis could not be rejected at the 0.05 level.
- In short, this approach is to specify the significance level in advance and use p-value only to determine whether or not the null hypothesis can be rejected at the stated significance level.
- Many practitioners find this approach to hypothesis testing not only too rigid, but basically illogical. It is very reasonable to have more confidence that the null hypothesis is false with a p-value of 0.0001 than with a p-value of 0.042?

p-values

- The less likely the obtained results (or more extreme results) under the null hypothesis, the more confident one should be that the null hypothesis is false.
- The null hypothesis should not be rejected once and for all. The possibility that it was falsely rejected is always present, and, all else being equal, the lower the p-value, the lower this possibility.
- According to this view, research reports should not contain the p-value, only whether or not the values were significant (at or below the significance level).
- However it is much more reasonable to just report the p-values. That way each reader can make up his or her mind about just how convinced they are that the null hypothesis is false.

0.6 Interpreting p-values

The p-value is the probability of having observed our data (or more extreme data) when the null hypothesis is true

The smaller the p-value, the less likely it is that the sample results come from a situation where the null hypothesis H_0 is true. If the p-value is sufficiently small, we reject the null hypothesis, and support the alternative hypothesis H_a .

One Sided Tests

- p-value > 0.05 : no evidence against H_0 in favour of H_a
- p-value < 0.05 : evidence against H_0 in favour of H_a

Two Sided Tests

- p-value > 0.025 : no evidence against H_0 in favour of H_a
- p-value < 0.025 : evidence against H_0 in favour of H_a

We will use a simplistic system for interpreting significance values (i.e. p-values).

If a p value is less than 0.02 we reject the null hypothesis. If the p value is greater than 0.05 we fail to reject the null hypothesis

If the pvalue is between the two thresholds then we deem the procedure to be inconclusive.

The null hypothesis is that the true correlation coefficient is zero (which is to say, no linear relationship exists).

0.7 Hypothesis Testing and p-values

0.7.1 p value

The P-value (or p-value or probability value is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming the null hypothesis is true. The null hypothesis is rejected if the P-value is very small, such as less than 0.05.

- In hypothesis tests, the difference between the observed value and the parameter value specified by H_0 is computed and the probability of obtaining a difference this large or large is calculated.
- The probability of obtaining data as extreme, or more extreme, than the expected value under the null hypothesis is called the **p-value**.
- There is often confusion about the precise meaning of the p-value probability computed in a significance test. It is not the probability of the null hypothesis itself.
- Thus, if the probability value is 0.0175, this does not mean that the probability that the null hypothesis is either true or false is 0.0175.
- It means that the probability of obtaining data as different or more different from the null hypothesis as those obtained in the experiment is 0.0175.

0.7.2 Using P-values to reject the null hypothesis

- According to one view of hypothesis testing, the significance level should be specified before any statistical calculations are performed. Then, when the p-value is computed from a significance test, it is compared with the significance level.
- The null hypothesis is rejected if p-value is at or below the significance level; it is not rejected if p-value is above the significance level. The degree to which p ends up being above or below the significance level does not matter.

0.7.3 P-values

- The null hypothesis either is or is not rejected at the previously stated significance level. Thus, if an experimenter originally stated that he or she was using the $\alpha = 0.05$ significance level and p-value was subsequently calculated to be 0.042, then the person would reject the null hypothesis at the 0.05 level.
- If p-value had been 0.0001 instead of 0.042 then the null hypothesis would still be rejected at the 0.05 significance level.
- The experimenter would not have any basis to be more confident that the null hypothesis was false with a p-value of 0.0001 than with a p-value of 0.041.
- Similarly, if the p had been 0.051 then the experimenter would fail to reject the null hypothesis

0.7.4 p-values

- He or she would have no more basis to doubt the validity of the null hypothesis than if p-value had been 0.482. The conclusion would be that the null hypothesis could not be rejected at the 0.05 level.
- In short, this approach is to specify the significance level in advance and use p-value only to determine whether or not the null hypothesis can be rejected at the stated significance level.
- Many statisticians and researchers find this approach to hypothesis testing not only too rigid, but basically illogical. It is very reasonable to have more confidence that the null hypothesis is false with a p-value of 0.0001 than with a p-value of 0.042?
- The less likely the obtained results (or more extreme results) under the null hypothesis, the more confident one should be that the null hypothesis is false. The null hypothesis should not be rejected once and for all. The possibility that it was falsely rejected is always present, and, all else being equal, the lower the p-value, the lower this possibility.
- According to this view, research reports should not contain the p-value, only whether or not the values were significant (at or below the significance level).
- However it is much more reasonable to just report the p-values. That way each reader can make up his or her mind about just how convinced they are that the null hypothesis is false.