Hypothesis Testing for Beginners

Michele Piffer

LSE

August, 2011

One year ago a friend asked me to put down some easy-to-read notes on hypothesis testing. Being a student of Osteopathy, he is unfamiliar with basic expressions like "random variables" or "probability density functions". Nevertheless, the profession expects him to know the basics of hypothesis testing.

These notes offer a very simplified explanation of the topic. The ambition is to get the ideas through the mind of someone whose knowledge of statistics is limited to the fact that a probability cannot be bigger than one. People familiar with the topic will find the approach too easy and not rigorous. But this is fine, these notes are not intended to them.

For comments, typos and mistakes please contact me on m.piffer@lse.ac.uk

Plan for these notes

- Describing a random variable
  - Expected value and variance
  - Probability density function
  - Normal distribution
  - Reading the table of the standard normal
- Hypothesis testing on the mean
  - The basic intuition
  - Level of significance, p-value and power of a test
  - An example

Random Variable: Definition

- The first step to understand hypothesis testing is to understand what we mean by "random variable"
- A *random variable* is a variable whose realization is determined by chance
- You can think of it as something intrinsically random, or as something that we don't understand completely and that we call "random"

Random Variable: Example 1

- What is the number that will show up when rolling a dice? We don't know what it is ex-ante (i.e. before we roll the dice). We only know that numbers from 1 to 6 are equally likely, and that other numbers are impossible.

- Of course, we would not consider it random if we could keep track of all the factors affecting the dice (the density of air, the precise shape and weight of the hand...). Being impossible, we refer to this event as random.

- In this case the random variable is {Number that appears when rolling a dice once} and the possible realizations are {1, 2, 3, 4, 5, 6 }

Random Variable: Example 2

- How many people are swimming in the lake of Lausanna at 4 pm? If you could count them it would be, say, 21 on June 1st, 27 on June 15th, 311 on July 1st...
- Again, we would not consider it random if we could keep track of all the factors leading people to swim (number of tourists in the area, temperature, weather..). This is not feasible, so we call it a random event.
- In this case the random variable is {Number of people swimming in the lake of Lausanne at 4 pm} and the possible realizations are {0, 1, 2, ... }

Moments of a Random Variable

- The world is stochastic, and this means that it is full of random variables. The challenge is to come up with methods for describing their behavior.
- One possible description of a random variable is offered by the *expected value*. It is defined as the sum of all possible realizations weighted by their probabilities
- In the example of the dice, the expected value is 3.5, which you obtain from

$$E[x] = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3,5$$

Moments of a Random Variable

- The expected value is similar to an average, but with an important difference: the average is something computed ex-post (when you already have different realizations), while the expected value is ex-ante (before you have realizations).
- Suppose you roll the dice 3 times and obtain $\{1, 3, 5\}$. In this case the average is 3, although the expected value is 3,5.
- The variable is random, so if you roll the dice again you will probably get different numbers. Suppose you roll the dice again 3 times and obtain $\{3, 4, 5\}$. Now the average is 4, but the expected value is still 3,5.
- The more observations extracted and the closer the average to the expected value

Moments of a Random Variable

- The expected value gives an idea of the most likely realization. We might also want a measure of the volatility of the random variable around its expected value. This is given by the *variance*
- The variance is computed as the expected value of the squares of the distances of each possible realization from the expected value
- In our example the variance is 2,9155. In fact,

x	probability	x-E(x)	(x-E(x))^2	prob*(x-E(x))^2
1	0,1666	-2,5	6,25	1,04125
2	0,1666	-1,5	2,25	0,37485
3	0,1666	-0,5	0,25	0,04165
4	0,1666	0,5	0,25	0,04165
5	0,1666	1,5	2,25	0,37485
6	0,1666	2,5	6,25	1,04125
				2,9155

Moments of a Random Variable

- Remember, our goal for the moment is to find possible tools for describing the behavior of a random variable
- Expected values and variances are two, fairly intuitive possible tools. They are called *moments* of a random variable
- More precisely, they are the first and the second moment of a random variable (there are many more, but we don't need them here)
- For our example above, we have
  - First moment $=$ E(X) $=$ 3,5
  - Second moment $=$ Var(X) $=$ 2,9155

PDF of a Random Variable

- The moments are only one possible tool used to describe the behavior of a random variable. The second tool is the probability density function

- A *probability density function* (pdf) is a function that covers an area representing the probability of realizations of the underlying values

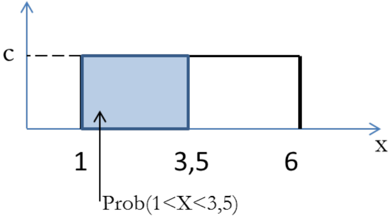
- Understanding a pdf is all we need to understand hypothesis testing

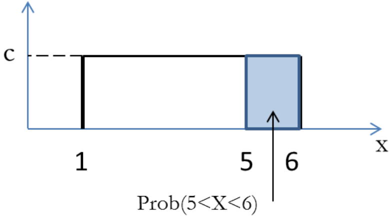
- Pdfs are more intuitive with continuous random variables instead of discrete ones (as from example 1 and 2 above). Let's move now to continuous variables

PDF of a Random Variable: Example 3

- Consider an extension of example 1: the realizations of X can still go from 1 to 6 with equal probability, but all intermediate values are possible as well, not only the integers $\{1, 2, 3, 4, 5, 6\}$

- Given this new (continuous) random variable, what is the probability that the realization of X will be between 1 and 3,5? What is the probability that it will be between 5 and 6? Graphically, they are the area under the pdf in the segments [1; 3,5] and [5; 6]



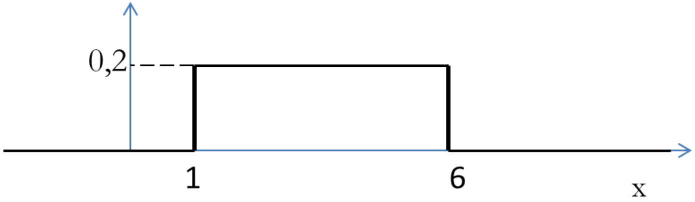


PDF of a Random Variable

- The graphic intuition is simple, but of course to compute these probabilities we need to know the value of the parameter c, i.e. how high is the pdf of this random variable
- To answer this question, you should first ask yourself a preliminary question: what is the probability that the realization will be between 1 and 6? Of course this probability is 1, as by construction the realization cannot be anywhere else. This means that the whole area under a pdf must be equal to one
- In our case this means that (6-1)*c must be equal to 1, which implies $c = 1/5 = 0{,}2$.

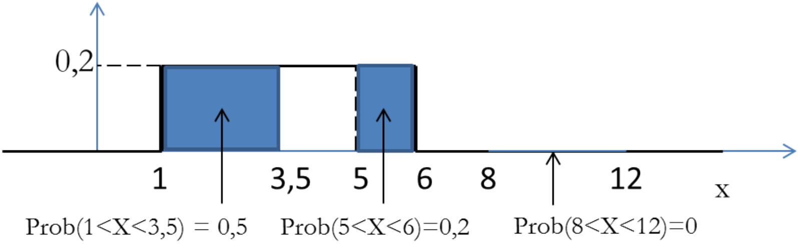
PDF of a Random Variable

- Now, ask yourself another question: what is the probability that the realization of X will be between 8 and 12? This probability is zero, given that [8, 12] is outside [1,6]
- This means that above the segment [8,12] (and everywhere outside the support [1, 6]) the area under the pdf must be zero.
- To sum up, the full pdf of this specific random variable is



PDF of a Random Variable

- In our example we see that:
  - $Prob(1 < X < 3,5) = (3,5 - 1) * 1/5 = 1/2$
  - $Prob(5 < X < 6) = (6 - 5) * 1/5 = 1/5$
  - $Prob(8 < X < 24) = (24 - 8) * 0 = 0$



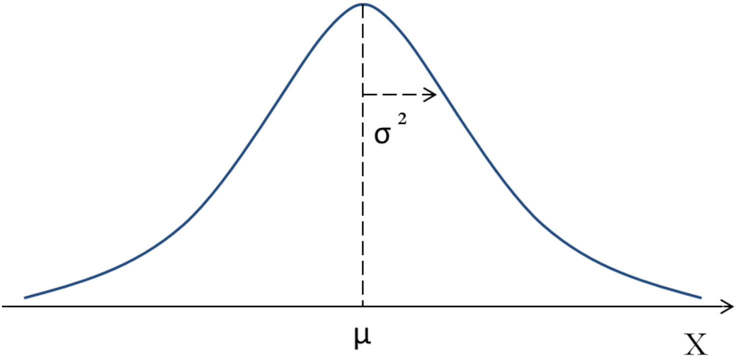
- Hypothesis testing will rely extensively on the idea that, having a pdf, one can compute the probability of all the corresponding events. Make sure you understand this point before going ahead

Normal Distribution

- We have seen that the pdf of a random variable synthesizes all the probabilities of realization of the underlying events
- Different random variables have different distributions, which imply different pdfs. For instance, the variable seen above is uniformly distributed in the support [1,6]. As for all uniform distributions, the pdf is simply a constant (in our case 0,2)
- Let's introduce now the most famous distribution, which we will use extensively when doing hypothesis testing: the normal distribution

Normal Distribution

- The *normal distribution* has the characteristic bell-shaped pdf:



Normal Distribution

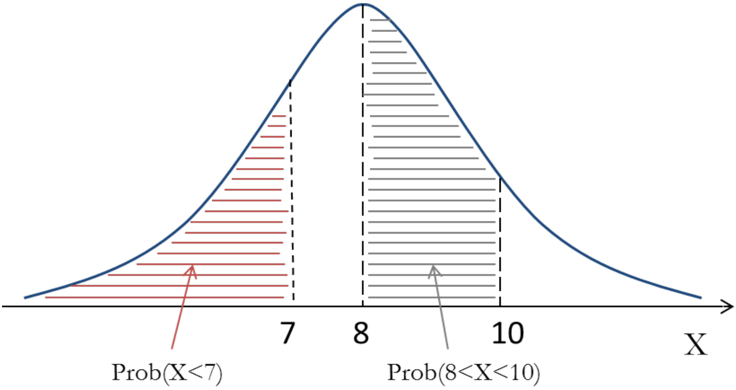
- Contrary to the uniform distribution, a normally distributed random variable can have realizations from $-\infty$ to $+\infty$, although realizations in the tail are really rare (the area in the tail is very small)
- The entire distribution is characterized by the first two moments of the variable: μ and σ^2. Having these two moments one can obtain the precise position of the pdf
- When a random variable is normally distributed with expected value μ and variance σ^2, then it is written

$$X \sim N(\mu, \sigma^2)$$

We will see this notation again

Normal Distribution

- Let's consider an example. Suppose we have a normally distributed random variable with $\mu = 8$ and $\sigma^2 = 16$. What is the probability that the realization will be below 7? What is the probability that it will be in the interval [8, 10]?



Standard Normal Distribution

- Graphically this is easy. But how do we compute them?
- To answer this question we first need to introduce a special case of the normal distribution: the standard normal distribution
- The *standard normal distribution* is the distribution of a normal variable with expected value equal to zero and variance equal to 1. It is expressed by the variable Z:

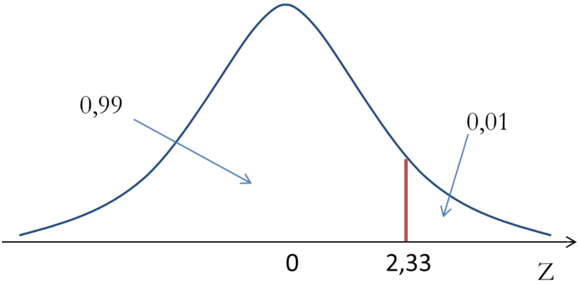
$$Z \sim N(0,1)$$

- The pdf of the standard normal looks identical to the pdf of the normal variable, except that it has center of gravity at zero and has a width that is adjusted to allow the variance to be one

Standard Normal Distribution

- What is the big deal of the standard normal distribution? It is the fact that we have a table showing, for each point in $[-\infty, +\infty]$, the probability that we have to the left and to the right of that point.

- For instance, one might want to know the probability that a standard normal has a realization lower that point 2.33. From these table one gets that it is 0,99. What is then the probability that the realization is above 2,33? 0,01, of course.



Standard Normal Distribution

- The following slide shows the table for the standard normal distribution. On the vertical and horizontal axes you read the z-point (respectively the tenth and the hundredth), while inside the box you read the corresponding probability

- The blue box shows the point we used to compute the probability in the previous slide

- Make sure you can answer the following questions before going ahead
  - Question 1: what is the probability that the standard normal will give a realization below 1?
  - Question 2: what is the point z below which the standard normal has 0,95 probability to occur?
  - Question 3: what is the probability that the standard normal will give a realization above (not below) -0.5

Tables of the Normal Distribution

Probability Content from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Standard Normal Distribution

- Answers to our previous questions
  - 1) $Prob(Z < 1) = 0,8413$
  - 2) z so that $Prob(Z < z) = 0,95$ is $1,64$
  - 3) $1 - Prob(Z < 0.5) = 0,6915$

- Last, make sure you understand the following theorem: if X is normally distributed with expected value μ and variance σ^2, then if you subtract μ from X and divide everything by $\sqrt{\sigma^2} = \sigma$ you obtain a new variable, which is a standard normal:

$$\text{Given } X \sim N(\mu, \sigma^2) \ \rightarrow \ \frac{X - \mu}{\sigma} \sim N(0,1) = Z$$

Normal Distribution

- We are finally able to answer our question from a few slides before: what is the value of $Prob(X < 7)$ and $Prob(8 < X < 10)$ given $X \sim N(8, 16)$?
- Do we have a table for this special normal variable with E(x)=8 and Var(x)=16? No! But we don't need it: we can do a transformation to X and exploit the table of the standard normal
- Remembering that you can add/subtract or multiply/divide both sides of an inequality, we obtain..

Normal Distribution

$$Prob(X < 7) = Prob(X - 8 < 7 - 8) = Prob(X - 8 < -1) =$$
$$= Prob(\frac{X - 8}{4} < -\frac{1}{4}) = Prob(\frac{X - 8}{4} < -0,25)$$

- The theorem tells us that the variable $\frac{X-\mu}{\sigma} = \frac{X-8}{4}$ is a standard normal, which greatly simplifies our question to: what is the probability that a standard normal has a realization on the left of point $-0,25$? We know how to answer this question. From the table we get

$$Prob(X < 7) = Prob(Z < -0,25) = 1 - Prob(Z > 0.25)$$
$$= 1 - 0,5987 = 0.4013 = 40,13\%$$

Normal Distribution

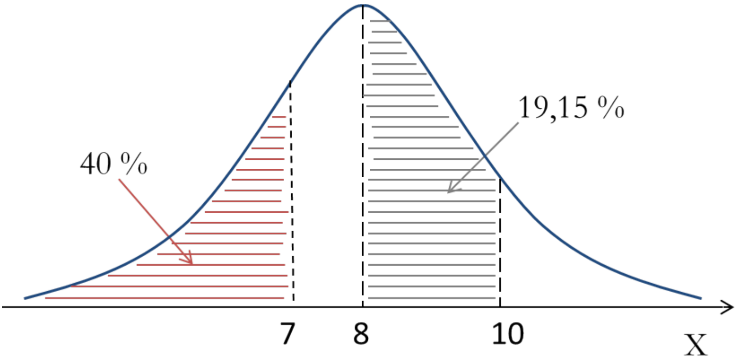
- Let's now compute the other probability:

$$
\begin{aligned}
Prob(8 < X < 10) &= Prob(X < 10) - Prob(X < 8) = \\
&= Prob(Z < \frac{10-8}{4}) - Prob(Z < \frac{8-8}{4}) \\
&= Prob(Z < 0.5) - Prob(Z < 0) = 0,6915 - 0.5 = \\
&= 0,1915 = 19,15\%
\end{aligned}
$$

- To sum up, a normal distribution with expected value 8 and variance 16 could have realizations from $-\infty$ to ∞. But we now know that the probability that X will be lower than 7 is 40,13 %, while the probability that it will be in the interval [8,10] is 19,15 %

Normal Distribution

- Graphically, given $X \sim N(8, 16)$



- Exercise: show that $Prob(X > 11) = 22,66\%$

Hypothesis Testing

- We are finally able to use what we have seen so far to do hypothesis testing. What is this about?

- In all the exercises above, we assumed to know the parameter values and investigated the properties of the distribution (i.e. we knew that $\mu = 8$ and $\sigma^2 = 16$). Of course, knowing the true values of the parameters is not a straightforward task

- By *inference* we mean research of the values of the parameters given some realizations of the variable (i.e. given some data)

Hypothesis Testing

- There are two ways to proceed. The first one is to use the data to estimate the parameters, the second is to guess a value for the parameters and ask the data whether this value is true. The former approach is <u>*estimation*</u>, the latter is <u>*hypothesis testing*</u>

- In the rest of the notes we will do hypothesis testing on the expected value of a normal distribution, assuming that the true variance is known. There are of course tests on the variance, as well as all kinds of tests

Hypothesis Testing

- Consider the following scenario. We have a normal random variable with expected value μ unknown and variance equal to 16
- We are told that the true value for μ is either 8 or 25. For some external reason we are inclined to think that $\mu = 8$, but of course we don't really know for sure.
- What we have is one realization of the random variable. The point is, what is this information telling us? Of course, the closer is the realization to 8 and the more I would be willing to conclude that the true value is 8. Similarly, the closer the realization is to 25 and the more I would reject that the hypothesis that $\mu = 8$

Hypothesis Testing

- How do we use the data to derive some meaningful procedure for inferring whether my data is supporting my null hypothesis of $\mu = 8$ or is rejecting it in favor of the alternative $\mu = 25$?

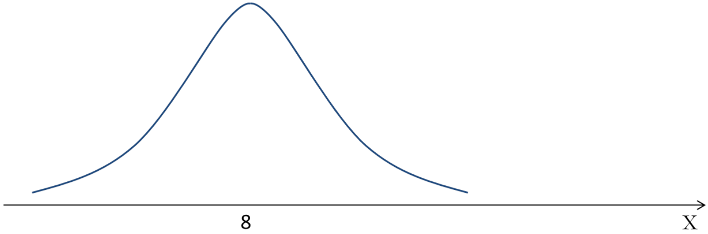
- Let's first put things formally:

$$X \sim N(\mu, 16) \quad \text{, with } H_0 : \mu = 8 \text{ vs. } H_a : \mu = 25$$

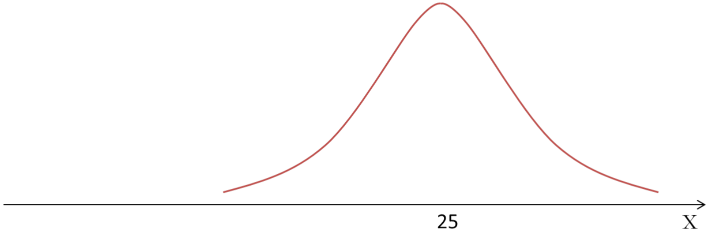
- Call x' the realization of X which we have and which we use to do inference. What is the distribution that has generated this point, $X \sim N(8, 16)$ or $X \sim N(25, 16)$?

Hypothesis Testing

- If null hypothesis is true, the realization x' comes from

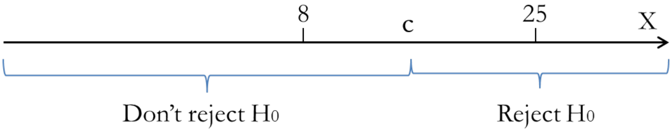


- If alternative hypothesis is true, the realization x' comes from



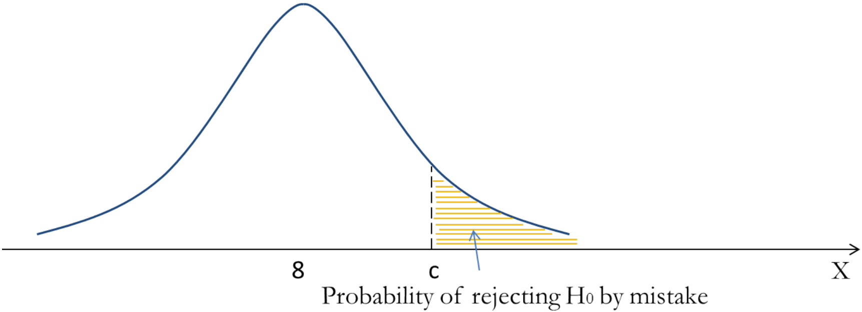
Hypothesis Testing

- The key intuition is: how far away from 8 must the realization be for us to reject the null hypothesis that the true value of μ is 8?
- Suppose $x' = -12$. In this case x' is much more likely to come from $X \sim N(8, 16)$ rather than from $X \sim N(25, 16)$. Similarly, if $x' = 55$ then this value is much more likely to come from $X \sim N(25, 16)$ rather than from $X \sim N(8, 16)$
- The procedure is to choose a point c (called *critical value*) and then reject H_0 if $x' > c$, while not reject H_0 if $x' < c$. The point is to find this c in a meaningful way



Hypothesis Testing

- Start with the following exercise. For a given c, what is the probability that we reject H_0 by mistake, i.e., when the true value for μ was actually 8? As we have seen, this is nothing but thee area under $X \sim N(8, 16)$ on the interval $[c, +\infty]$



Hypothesis Testing

- Given c, we also know how to compute this probability. Under H_0, it is simply $Prob(X > c) = Prob(Z > \frac{c-8}{4})$
- What we just did is: given c, compute the probability of rejecting H_0 by mistake. The problem is that we don't have c yet, that's what we need to compute!
- A sensible way to choosing c is to do the other way around: find c so that the probability of rejecting H_0 by mistake is a given number, which we choose at the beginning. In other words, you choose with which probability of rejecting H_0 by mistake you want to work and then you compute the corresponding critical value

Hypothesis Testing

- In jargon, the probability of rejecting H_0 by mistake is called "type 1 error", or *level of significance*. It is expressed with the letter α
- Our problem of finding a meaningful critical value c has been solved: at first, choose a confidence interval. Then find the point c corresponding to the interval α. Having c, one only needs to compare it to the realization x' from the data. If $x' > c$, then we reject $H_0 = 8$, knowing that we might be wrong in doing so with probability α
- The most frequent values used for α are 0.01, 0,05 and 0,1

Hypothesis Testing

- Let's do this with our example. Suppose we want to work with a level of significance of 1 %

- The first thing to do is to find the critical vale c which, if H_0 was true, would have 1 % probability on the right. Formally, find c so that

$$0,01 = Prob(X > c) = Prob(Z > \frac{c-8}{4})$$

- What is the point of the standard normal distribution which has area of 0,01 on the right? This is the first value we saw when we introduced the N(0,1) distribution. The value was 2,33:

$$0,01 = Prob(Z > \frac{c-8}{4}) = Prob(Z > 2,33)$$

Impose $\frac{c-8}{4} = 2,33$ and obtain $c = 17,32$.

Hypothesis Testing

- At this point we have all we need to run our test. Suppose that the realization from X that we get is 15. Is x'=15 far enough from $\mu = 8$ to lead us to reject H_0? Clearly $15 < 17,32$, which means that we cannot reject the hypothesis that $\mu = 8$
- Suppose we run the test with another realization of X, say $x' = 20$. Is 20 sufficiently close to $\mu = 25$ to lead us to reject H_0? Of course it is, given that 20 is above the critical $c = 17,32$

Hypothesis Testing

- Note that when rejecting H_0 we cannot be sure at 100% that the true value for the parameter μ is not 8. Similarly, when we fail to reject H_0 we cannot be sure at 100 % that the true value for μ is actually 8.
- What we know is only that, in doing the test several times, we would mistakenly reject $\mu = 8$ with $\alpha = 0.01$ probability.

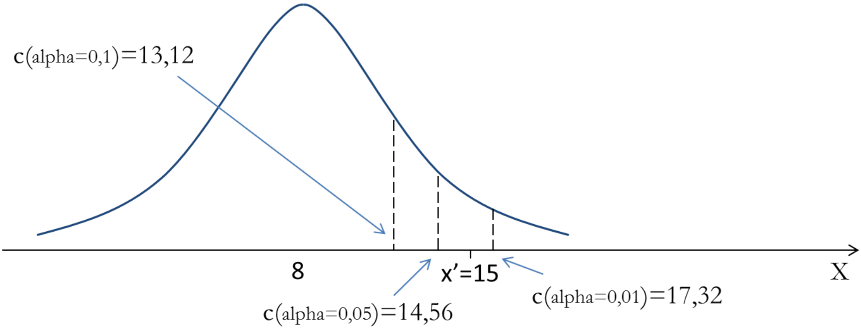
Hypothesis Testing

- Our critical value c would have been different, had we chosen to work with a different level of significance. For instance, the critical value for $\alpha = 0,05$ is 14,56, while the critical value for $\alpha = 0,1$ is 13,12 (make sure you know how to compute these values)

α	c
0.01	17.32
0.05	14.56
0.10	13.12

- Suppose that the realization of X is 15. Do we reject H_0 when working with a level of significance of 1 %? No. But do we reject it if we work with a level of significance of 10 %? Yes! The higher the probability that we accept in rejecting H_0 by mistake and the more likely it is that we reject H_0

Hypothesis Testing



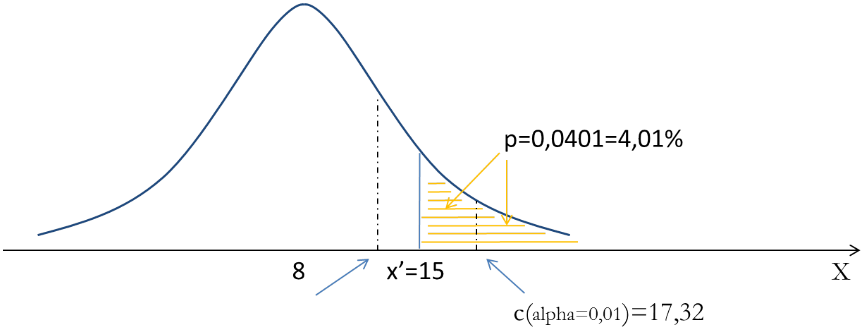
Hypothesis Testing

- If, given x', the outcome of the test depends on the level of significance chosen, can we find a probability p so that we will reject H_0 for all levels of significance from 100 % down to p?

- To see why this question is interesting, compute the probability p on the right of $x' = 15$. We will be able to reject H_0 for all critical values lower than x', or equivalently, for all levels of significance higher than p

$$p = Prob(X > x') = Prob(Z > \frac{15 - 8}{4}) = Prob(Z > 1,75) =$$
$$= 1 - Prob(Z < 1,75) = 0,0401$$

Hypothesis Testing



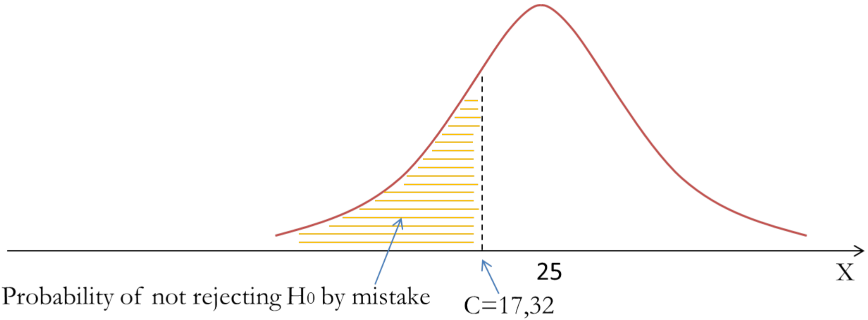
Hypothesis Testing

- If we have chosen $\alpha < p$, then it means that $x' < c$ and we cannot reject H_0
- If we have chosen an $\alpha > p$, then it means that $x' > c$ and we can reject H_0
- In our example, we can reject H_0 for all levels of significance $\alpha \geq 4,01\%$
- The probability p is called *p-value*. The p-value is the lowest level of significance which allows to reject H_0. The lower p and the easier will be to reject H_0, since we will be able to reject H_0 with lower probabilities of rejecting it by mistake. In other words, the lower p and the more confident you are in rejecting H_0

Hypothesis Testing

- Our last step involves the following question. What if the true value for μ was actually 25? What is the probability that we mistakenly fail to reject H_0 when the true μ was actually 25?

- Given what we saw so far this should be easy: it is the area on the region $[-\infty, c]$ under the pdf of the alternative hypothesis:



Hypothesis Testing

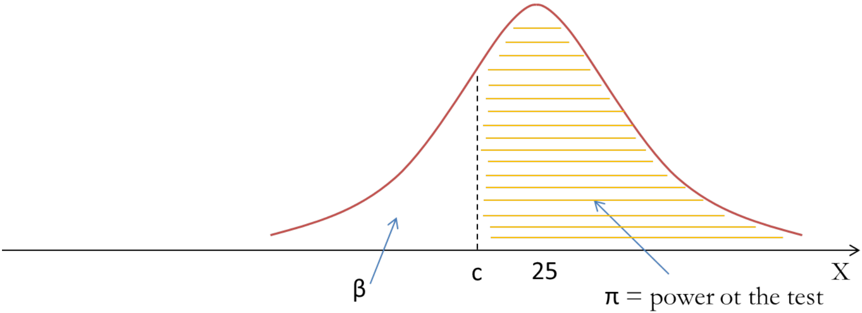
- In jargon, the probability of not rejecting H_0 when it is actually wrong is called "type 2 error". It is expressed by the letter β
- Let's go back to our example. At this point you should be able to compute β easily. Under H_a we have

$$\beta = Prob(X > c) = Prob(X > 17,32) =$$
$$= Prob(\frac{X-25}{4} < \frac{17,32-25}{4}) = Prob(Z < -1,92) = 0,0274$$

- Last, what is the probability that we reject H_0 when it was actually wrong and the true one was H_a? It is obviously the area under $X \sim N(25,16)$ in the interval $[c,+\infty]$, or equivalently, $1-\beta = 0,9726$. This is the power of the test

Hypothesis Testing

- The *power* of a test is the probability that we correctly reject the null hypothesis. It is expressed by the letter π

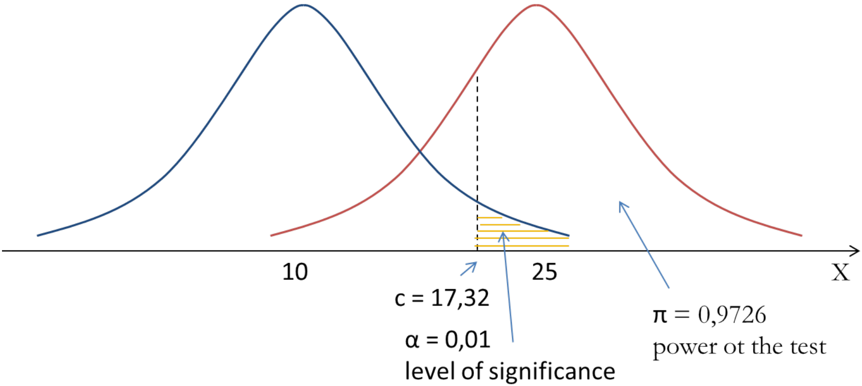


Hypothesis Testing

- To sum up, the procedure for running an hypothesis testing requires to:
  - Choose a level of significance α
  - Compute the corresponding critical value and define the critical region for rejecting or not rejecting the null hypothesis
  - Obtain the realization of your random variable and compare it to c
  - If you reject H_0, then you know that the probability that you are wrong in doing so is α
  - Given H_1 and the level c, compute the power of the test
  - If you reject the H_0 in favor of H_1, then you know that you are right in doing so with probability given by π
- Alternatively, one can compute the p-value corresponding to the realization of X given by the data and infer all levels of significance up to which you can reject H_0

Hypothesis Testing

- In the example seen so far, choosing $\alpha = 0,01$, we reject H_0 for all $x' > 17,32$ and do not reject it for all $x' < 17,32$



Hypothesis Testing

- To conclude, three remarks are necessary:
- Remark 1: What if we specify that the alternative hypothesis is not a single point, but it is simply $H_a : \mu \neq 8$?
- Such a test is called two-sided test, compared to the one-sided test seen so far. Nothing much changes, other than the fact that the rejection region will be divided into three parts. Find $c_1.c_2$ so that

$$\alpha = Prob(X < c_1 \text{ or } X > c_2) = 1 - Prob(c_1 < X < c_2)$$

- Then, you reject H_0 if x' lies outside the interval $[c_1, c_2]$ and do not reject H_0 otherwise

Hypothesis Testing

- Remark 2: why should one use only a single realization of X? Doesn't it make more sense to extract n observations and then compute the average?

- Yes, of course. Nothing much changes in the interpretation. The only note of caution is on the distribution used. It is possible to show that, if $X \sim N(\mu, \sigma^2)$, then

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

  with $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

- The construction of the test follows the same steps, but you have to use the new distribution

Hypothesis Testing

- Remark 3: so far we assumed that we knew the true value of σ^2. But what is it? Shouldn't we estimate it? Doesn't this change something?

- Yes, or course. One can show that the best thing to do is to substitute σ^2 with the following statistic:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_1 - \overline{X})^2$$

- The important difference in this case is that the statistics will not follow the normal distribution, but a different distribution (known as the t distribution). The logical steps for constructing the test are the same. For the details, have a look at a book of statistics