**07 - Etapas de um teste**

[0:00] Pessoal, maravilha. Vamos agora ver as etapas de um teste de hipóteses. Abrindo o nosso Colab no tópico "2.2 Etapas Básicas de um Teste" Encontraremos algumas etapas, que são as que eu escolhi para seguirmos em todos os testes que vamos executar.

[0:09] No primeiro teste, nós não fizemos, mas nos próximos a gente vai seguir essa sequência de cinco etapas até conseguirmos finalizar o teste. Depois que você entender todo o funcionamento, eu ensinarei uma forma simples de executar esse teste de hipóteses e tomar decisões.

[0:29] Mas, para isso, você tem que entender o que está acontecendo, que hipótese você está testando, que distribuição eu estou utilizando. Eu posso precisar trocar essa distribuição.

[0:39] Existe um parâmetro dentro do método do Python que eu preciso modificar. E se eu não entender o que está acontecendo, eu não sei o que eu troco ali dentro.

[0:50] Outra motivação de estudarmos o teste em etapas é, por exemplo: tem alguém realizando o curso que precisa fazer uma prova escrita e o professor quer que você realmente siga essas etapas.

[1:03] Então, aqui você vai aprender a realizar todas essas etapas. Lógico, aqui usando o computador, na prova você vai ter que fazer os cálculos manualmente. Mas, não se preocupe, não é nada muito complicado.

[1:15] Vamos começar com a primeira etapa, que é a formulação das hipóteses. Eu não vou ser tão extenso neste vídeo, porque a gente vai continuar esse assunto nos outros vídeos, então eu não preciso ficar falando muito.

[1:27] Mas eu deixei esse conteúdo como um material de estudo para você, de consulta também, você encontra tudo pronto no notebook, assim saberá como proceder e o que significa cada uma dessas etapas.

[1:38] Sobre a primeira formulação das hipóteses, como eu disse, nós sempre temos duas hipóteses: um teste, que é o H0, hipótese nula, e H1 que é hipótese alternativa.

[1:48] O H0 sempre vai ter a igualdade no teste. Está escrito aqui: a hipótese nula sempre afirma uma igualdade ou propriedade populacional. E H1 é ao contrário, é a negação disso.

[2:00] Quando eu digo igualdade, significa: igual, simples, ou então maior/igual, ou menor/igual. Isso tudo vai estar no H0, lembre-se disso.

[2:10] O H1, ao contrário, vai ser diferente: o menor e o maior simples. A gente vai ver como formar essas hipóteses todas ao longo do treinamento.

[2:22] O segundo passo é: a escolha da distribuição amostral adequada. Já começamos a estudar, mas não nos aprofundamos muito. Eu construí um esquema que vai nos ajudar em todas essas etapas.

| **n é igual ou maior a 30?** |  |  |  |
| --- | --- | --- | --- |
| Para reposta **"sim"** | O desvio padrão (sigma) é conhecido? | **Se sim**, então a média é mais ou menos igual a z vezes s sobre raiz de n | **Se não**, então a média é mais ou menos igual a z vezes o desvio padrão sobre raiz de n |
| Para resposta **"não"** | É possível afirmar que a população se distribui como uma normal? | **Se sim**, o desvio padrão é conhecido? **Sim:** a média é mais ou menos z vezes o desvio padrão sobre raiz de n; **Não:** média é mais ou menos t vezes s sobre raiz de n | **Se não**, aumente o tamanho da amostra ou utilize testes não-paramétricos |

[2:34] Esse esquema sempre nos auxiliará nas nossas escolhas. A primeira, é uma regra de decisão.

[2:44] Se o N for maior ou igual a 30, você se lembra do teorema do limite central, que a gente fez uma "provinha", uma coisa até bonita, para a gente entender o funcionamento disso e porquê do teorema do limite central.

[2:58] Aqui, a mesma coisa. Se o N for maior ou igual a 30, a gente tem uma opção, um caminho a seguir. Se o Sigma é conhecido – você sabe o que é o Sigma, o desvio padrão – a gente tem um caminho a ser seguido.

[3:11] Se não, a gente tem outro caminho. No caso, a gente está usando o Z, da distribuição normal. Se ele é conhecido, a gente usa ele mesmo; se ele não é conhecido, a gente usa a estimativa da amostra para o desvio padrão.

[3:24] Aqui, caso o N não seja maior que zero, a gente tem uma primeira pergunta: é possível afirmar que a população segue uma distribuição normal? Se sim, a gente segue por este caminho e novamente pergunta: o Sigma é conhecido?

[3:37] Se for, mantemos isso tudo utilizando a distribuição normal e utilizamos Sigma. Se não, a gente vai conhecer logo, logo a distribuição T de Student. Eu vou mostrar para você a tabela como consultar. E a gente usa o S também.

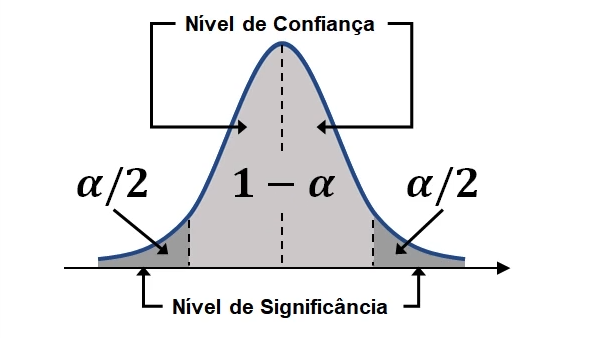
[3:54] Caso contrário, isto é, se isso não for uma distribuição normal, a gente ou aumenta o N, ou então a gente podemos partir para testes não paramétricos também. Tem essa opção. A gente vai estudar testes não paramétricos também.

[4:04] Então, vamos lá. Na terceira etapa, a gente fixa a significância do teste, o Alfa, que vai determinar as áreas de aceitação e de rejeição do teste.

[4:15] Os valores mais frequentes, como já aprendemos nos outros cursos são: 10%, 5% e 1%, o que gera um nível de confiança de 90%, 95% e 99%, respectivamente.

[4:29] Aqui está aquela figura do nível de confiança e do nível de significância, para um teste bicaudal. Já vamos aprender o que é bicaudal. Dentro dela há o nível de confiança: 1 menos Alfa.

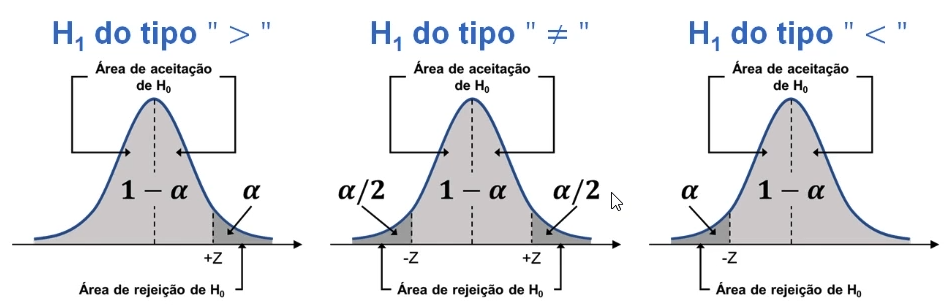
[4:43] E as caudas têm o valor de Alfa dividido por dois, porque elas são idênticas. Lembre-se de que isso é uma função simétrica, eu tenho Alfa sobre dois de um lado e Alfa sobre dois do outro.



[4:57] A gente já conhece bem o que é nível de confiança e significância. Depois temos os tipos de testes que vamos aplicar. O primeiro, H1, é hipótese alternativa. Se ela for do tipo maior, eu vou ter um teste bicaudal.

[5:09] Por que caudal? As duas pontas aqui das distribuições são conhecidas como caudas da distribuição, por isso bicaudal e unicaudal. Eu vou ter também um teste unicaudal superior, quando essa condição existir aqui nas nossas hipóteses.

[5:25] Quando na hipótese o H1 for diferente, o H0 vai ser igual, a gente tem um teste bicaudal. E quando o H1 tiver a condição menor, a gente tem um unicaudal inferior.



[5:43] Vamos estudar tudo isso nos testes que vamos executar. Lembrando, também, que é bom, para você fixar tudo isso, fazer os exercícios. Tem que fazer exercícios, não tem como fugir.

[5:58] O próximo passo é: calculo da estatística de teste. Você deve lembrar, a gente fez a normal, no vídeo anterior, e ele soltou dois parâmetros. Um é a estatística teste que eu falei e o outro é o P valor.

[6:09] Justamente, o primeiro parâmetro é ele, o cálculo de estatística de teste. Com ele definido, eu consigo ver se o parâmetro caiu na área de aceitação, de H0 - e já está marcada a área de aceitação nos nossos gráficos - ou na área de rejeição de H0. Perfeito?

[6:23] Isso vai ajudar a gente a tomar a decisão do teste, rejeitar H0 ou aceitar H0 e por aí vai. O quinto, e último, passo é justamente esse: rejeitar a hipótese nula ou aceitar a hipótese nula, e nós vamos aprender as regras de decisão.

[6:42] Vai ter uma figurinha legal para cada um deles. Para cada tipo de teste eu criei uma figura já, com a estatística de teste que você usa, as regras para rejeitar H0, tanto utilizando o valor crítico quanto o P valor.

[6:54] O valor crítico seria utilizar essa estatística de teste.

[7:01] Esses são os passos básicos. A gente vai acompanhar cada um dos testes seguindo todos esses passos e vamos aprender como funciona cada um deles, com um probleminha prático que pode acontecer no dia a dia.

[7:13] Então, no próximo vídeo, seguiremos para o início de testes paramétricos, beleza? Até lá.