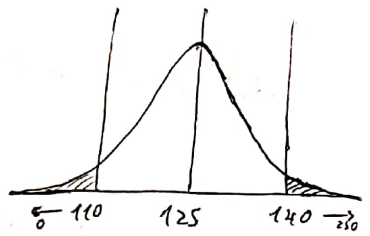
Guillermo Abad López

TME - Part 2

1)

a)

$$p(x \geq 140 \text{ or } x \leq 110) = \sum_{i=140}^{250} p(i) + \sum_{i=0}^{110} p(i)$$



where $p(i) = \frac{250!}{(250-i)!\, i!} 0{,}5^i \, 0{,}5^{250-i}$ is the probability distribution of obtaining i nº of heads,

if the coin is fair $p(heads) = 0{,}5 = p(tails)$, and because, the pd. is symmetric, we have:

$$p(x \geq 140 \text{ or } x \leq 110) = 2 \cdot \sum_{i=140}^{250} \frac{250!}{(250-i)!\, i!} 0{,}5^i \, 0{,}5^{250-i} = 2 \cdot 0{,}03321057 \simeq 0{,}0664 \rightarrow \boxed{6{,}64\,\%}$$

So, we see that the 7% arises from considering the probabilities of all the cases that would be worse, and adding them up.

The term that describes this value in the theory of NHST is the p-value:

$$p\text{-value}(K) = P(x = K \text{ or worse} \mid H_0)$$ where H_0 is the null hypotesis, in this case would be, that the coin is fair.

Finally say that this corresponds to a two-sided test, because we are taking into account both tails, that heads is more propense, or that tails is more propense, visually it's even easier to see, from the above graph, we are taking into account both sides.

b)

For rejecting H_0 we would need $p < \alpha$, because p can be interpreted as:

p-value: Probability that H_0 is correct from what we have just observed

and α, can be interpreted, as which of those probabilities will make you reject H_0, so:

$\alpha \uparrow$ you will reject H_0 easily (open-minded)

$\alpha \downarrow$ you won't reject H_0 easily (conservative)

Finally comming back to our exercice, for

$\alpha = 0{,}1$ we would reject H_0, and would think the coin is not fair.

$\alpha = 0{,}05$ we wouldn't reject H_0, and would still hold to the null hypotesis that the coin is fair, we are more conservative.

c)

To reject H_0 at a significance of $\alpha = 0{,}01$, we would need:

$2 \cdot p(x \geq A) < 0{,}01$ where A is the number of heads necessary to reject H_0

$$2 \cdot \sum_{i=A}^{250} \frac{250!}{(250-i)!\, i!} 0{,}5^{i} \cdot 0{,}5^{250-i} < 0{,}01 \longrightarrow A \geq 146$$

So we would need 146 heads to reject at $\alpha = 0{,}01$. (The p-value would be $p = 0{,}0094$)

d)

i)

Fix $\alpha = 0{,}05$, and compute the power of $\begin{cases} H_1: & p = 0{,}55 \\ H_2: & p = 0{,}6 \end{cases}$

Because power $= P(\text{Reject } H_0 \mid H_A)$, we first need to find when we reject H_0 for a significance $\alpha = 0{,}05$. If we make the same as before:

$2 \cdot p(x \geq A) < 0{,}05 \longrightarrow A \geq 141$ (with p-value $= 0{,}0497$ for $A = 141$)

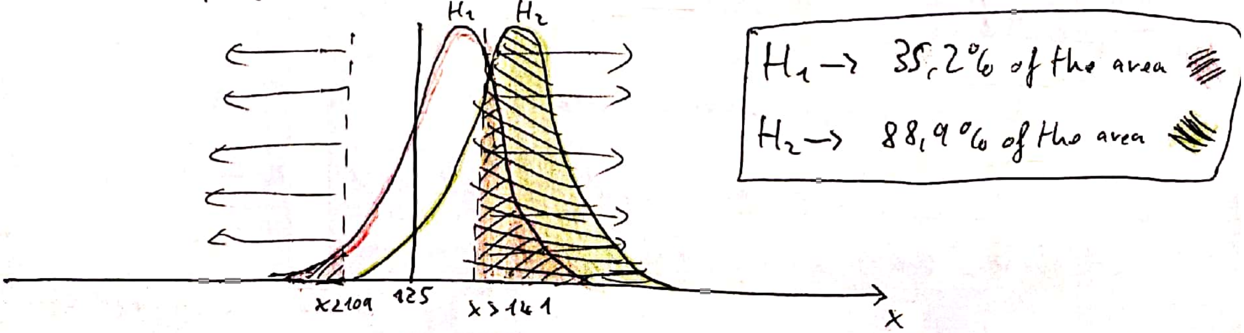
So for rejecting H_0 we need 141 heads or tails. Let's compute the powers:

$$\boxed{\text{power}_1} = P(x \geq 141 \text{ or } x \leq 109 \mid H_1) = \sum_{i=141}^{250} p(i \mid H_1) + \sum_{i=0}^{109} p(i \mid H_1) =$$
$$= \sum_{i=141}^{250} \frac{250!}{(250-i)!\, i!} 0{,}45^{i}\, 0{,}55^{250-i} + \sum_{i=0}^{109} \frac{250!}{(250-i)!\, i!} 0{,}45^{i}\, 0{,}55^{250-i} = \boxed{0{,}3524}$$

$$\boxed{\text{power}_2} = P(x \geq 141 \text{ or } x \leq 109 \mid H_2) = \sum_{i=141}^{250} p(i \mid H_2) + \sum_{i=0}^{109} p(i \mid H_2) =$$
$$= \sum_{i=141}^{250} \frac{250!}{\cdots} 0{,}6^{i}\, 0{,}4^{250-i} + \sum_{i=1}^{109} \frac{250!}{\cdots} 0{,}6^{i}\, 0{,}4^{250-i} = \boxed{0{,}8896}$$

Where we see that power$_2$ > power$_1$, it's obvious that it's easier for us to reject the coin being fair, if it actually gives 60–40% rather than if it gives 55–45%.

ii)



We see that the rejection zone, occupies much more area of the p.d. in the H_2 hypotesis $\rightarrow$ power$_2$ ↑↑

e)

i)

We need a set of innequations to be fullfilled:

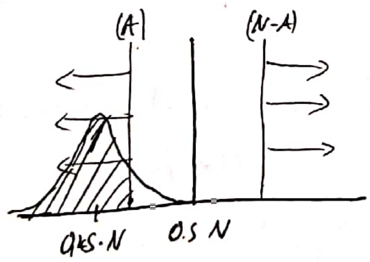
$$\left.\begin{aligned} p(x \geq A \text{ or } x \leq N-A \mid H_0) &\leq 0{,}05 \\ p(x \geq A \text{ or } x \leq N-A \mid H_A) &\geq 0{,}9 \end{aligned}\right\} \longrightarrow \boxed{N, A \text{ solutions}}$$

This set of innequalities can be writton exactly as:

$$\left.\begin{aligned} \sum_{i=N-A}^{N} \frac{N!}{(N-i)!\, i!} 0{,}5^i\, 0{,}5^{N-i} + \sum_{i=0}^{A} \frac{N!}{(N-i)!\, i!} 0{,}5^i\, 0{,}5^{N-i} &\leq 0{,}05 \\ \sum_{i=N-A}^{N} \frac{N!}{(N-i)!\, i!} 0{,}45^i\, 0{,}55^{N-i} + \sum_{i=0}^{A} \frac{N!}{(N-i)!\, i!} 0{,}45^i\, 0{,}55^{N-i} &\geq 0{,}9 \end{aligned}\right\} \rightarrow N, A \text{ solutions}$$

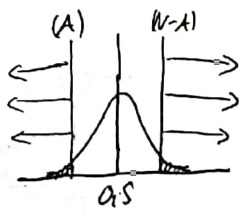
I couldn't solve this innequalities exactly, I even tried some RegionPlot[ ] in Mathematica, to annalize the phase-space $\{A, N\}$, but couldn't make it work at the end. So because this, I had to make it manually, trying values.

But I didn't pick blindly, because I have $p = 0{,}45$ ill have the most contribution for the sum of the power (2nd innequality), in $0{,}45 \cdot N$. And because I'll make a symmetric domain from $0 \to A$ and from $N-A \to N$, centered at $0{,}5 \cdot N$.



$> 0{,}9$

power summ
(2nd innequality)



$< 0{,}05$

p-value summ
(1st innequality)

I can pick a random N, then find the minimum A that makes power $> 0{,}9$ for that N, and check if the p-value for that $A < 0{,}05$, if it wasn't, there are no more solutions for that N, because that p-value is the minimum that satisfies power $\geq 0{,}9$. So I reapeted the process for bigger and bigger N's.

1) Pick N

2) With Mathematica find minimum A that fullfills power $\geq 0{,}9$

3) Check if that A fullfills p-value $< 0{,}05$.

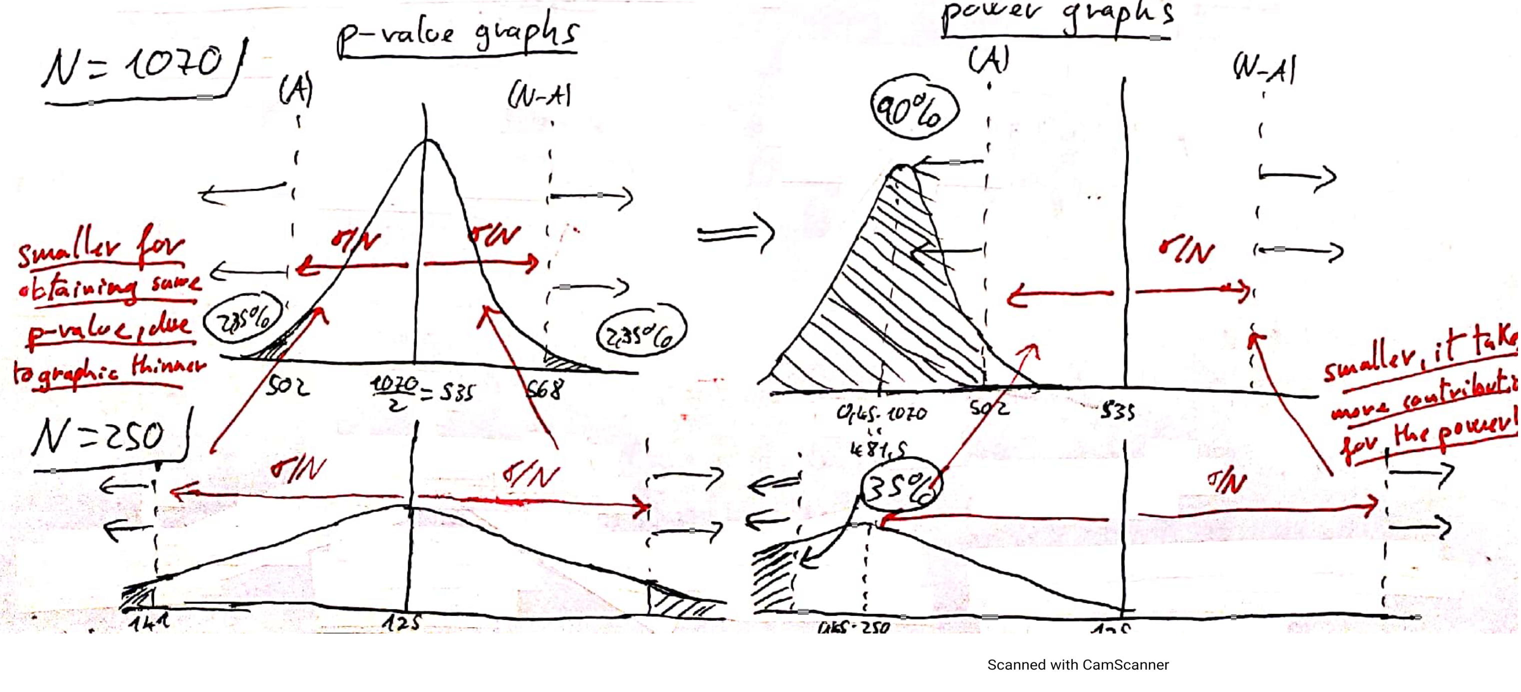
No → Repeat with bigger N

Yes → Repeat with the average of this N and the previous one that didn't had a solution.

I found a quite good solution for around $\boxed{N \approx 1070}$ $(A = 502)$

- $$\sum_{i=568}^{1070} \frac{1070!}{(1070-i)!\,i!}\,0{,}5^i\,0{,}5^{1070-i} + \sum_{i=0}^{502} \frac{1070}{(1070-i)!\,i!}\,0{,}5^i\,0{,}5^{1070-i} = \boxed{\begin{array}{c}\text{p-value:}\\ 0{,}047 \leq 0{,}05\end{array}} \checkmark\checkmark$$
- $$\sum_{i=568}^{1070} \frac{1070!}{(1070-i)!\,i!}\,0{,}45^i\,0{,}55^{1070-i} + \sum_{i=0}^{502} \frac{1070}{(1070-i)!\,i!}\,0{,}45^i\,0{,}55^{1070-i} = \boxed{\begin{array}{c}\text{power:}\\ 0{,}901 \geq 0{,}9\end{array}} \checkmark\checkmark$$

ii) At the end increasing N we achived that $\sigma = \frac{N}{2} - A$ that fullfills p-value $< 0{,}05$, is reduced enough to contribute more in the power.



f)

$$\boxed{P(H_A|data)_{50/50}} = \frac{P(data|H_A)\,P(H_A)}{P(data)} = \frac{\overbrace{P(data|H_A)}^{\circledast}\,P(H_A)}{\underbrace{P(data|H_0)}_{\text{from a)}}\,P(H_0) + \underbrace{P(data|H_A)}_{\circledast}\,P(H_A)} =$$

$$= \frac{0{,}4 \cdot \cancel{0{,}5}}{0{,}0664 \cdot \cancel{0{,}5} + 0{,}4 \cdot \cancel{0{,}5}} = \frac{0{,}4}{0{,}4664} = \boxed{0{,}857 = 86\%}$$

($\circledast$ The only Prob we didn't know was $P(data|H_A)$, which is:

$$P(x \geq 140 \text{ or } x \leq 110 \mid H_A) = \sum_{i=140}^{250} \frac{250!}{(250-i)!\,i!}\,0{,}45^i\,0{,}55^{250-i} + \sum_{i=0}^{110} \frac{250!}{(250-i)!\,i!}\,0{,}45^i\,0{,}55^{250-i} = 0{,}4$$)

g)

We have seen that if we consider a 50/50 chance for the dice fair or at H_A, we get results that benefit H_A.

If we divide $140/250 = 0{,}56$, so if we took $p = 0{,}56$ and $q = 0{,}44$ we would get a $p(H_B|data) = 0{,}888 \simeq 89\%$, which benefits H_B too.

But all these probabilities is considering there are the same H_0 and H_A or H_B coins around the world, which is not the case. If we consider that there are 1 of every 100 coins which has H_A and the rest has H_0 (is fair), which still wouldn't be near the reality. How many tricked coins have you seen? And how many fair ones? So, even with 1/99 proportion we get:

$$p(H_B|data)_{1/99} = \frac{0{,}5265 \cdot 0{,}01}{0{,}0664 \cdot 0{,}99 + 0{,}5265 \cdot 0{,}01} = \boxed{0{,}074 \simeq 7{,}4\% \text{ only}}$$

So, if we consider, that the vast majority of coins are fair, I would say the coin is most possible fair if we are in a normal enviorement.

If we are betting in some clandestin store, with the mafia, then maybe the prior probabilities are more near to 50/50 than 1/100, and the coin would be most certainly not fair. Depends on the situation!

2)

a)

In general, we have:

Type I error $\longleftrightarrow$ rejecting $H_0 \mid H_0 = \text{true}$ (false negative)

type II error $\longleftrightarrow$ not rejecting $H_0 \mid H_A = \text{true}$ (false positive)

in this case: $\begin{cases} H_0 \longrightarrow \text{is telling the truth} \\ H_A \longrightarrow \text{is lying} \end{cases}$, so we have:

$\boxed{\text{type I error: Tester thinks is lying, when it's telling the truth ("false lying")}}$

$\boxed{\text{type II error: Tester thinks is telling the truth, when it's lying ("false truth")}}$

and their probabilities from the table are:

	Testee tells truth	Testee lies
Tester thinks is telling truth	131	$\boxed{15}$ "false truth" type II error
Tester thinks is lying	9 "false lying" type I error	125

so:

$$\boxed{P(\text{type I error})} = \underset{(P(\text{reject truth} \mid \text{truth correct}))}{P(\text{reject } H_0 \mid H_0)} = \frac{9}{140} = 0.064 \approx \boxed{6.4\%}$$

$$\boxed{P(\text{type II error})} = \underset{(P(\text{not reject truth} \mid \text{lying correct}))}{P(\text{not reject } H_0 \mid H_A)} = \frac{15}{140} = 0.107 \approx \boxed{10.7\%}$$

b)

In NHST we have:

significance level: $\boxed{\alpha = P(\text{type I error})} = 0.064$

power: $\boxed{pow = 1 - P(\text{type II error})} = 1 - 0.107 = 0.893$

$(power = P(\text{rejecting } H_0 \mid H_A) = 1 - P(\text{not rejecting } H_0 \mid H_A))$

3)

I'm going to do this excercice in two ways, first I'll use an already done t-test from a python library, and then I'll do one myself with mathematica:

3.1)

In this case I'm going to use scipy.stats.ttest_1samp(data, 10); and repeat the experiment 1000 times to obtain a distribution of p-values, from where I will look to their mean.

a) Using "two sided" command for the test, we get:

$$\boxed{\text{p-value}_{\mu\neq 10} = 0{,}15 > \alpha} \longrightarrow \underline{\text{We do not reject } H_0}$$

b) Using "greater" this time, we get:

$$\boxed{\text{p-value}_{\mu>10} = 0{,}085 > \alpha} \longrightarrow$$

We do not reject H_0, but we are very close, for $\mu > 9$ we would get p-value $< \alpha$ and therefore we would reject H_0.

(I'll attach the distributions of the p-values I did obtain with python later)

3.2)

Now let's do the test our selfs, let's start by defining:

$$T = \frac{\hat{x} - \mu_0}{s/\sqrt{n}} \quad \text{where} \begin{cases} \hat{x} \text{ and } s \text{ are the mean and variance of our data} \\ \underset{(10)}{\mu_0} \text{ is the mean we want to test (mean of } H_0) \end{cases}$$

Thist T will be disributed following a Student's t distribution with n-1 dof:

$$p(t\,|\,H_0) = P_{t(n-1)}(t) = \frac{\Gamma(\frac{n}{2})\left(1+\frac{t^2}{n-1}\right)^{-n/2}}{\sqrt{(n-1)\pi}\;\Gamma(\frac{n-1}{2})}$$

(I've done 500 different sets of data with it's corresponding T, and then I have integrated each, and made an average, I'll also attach some code)

a)
$$\boxed{\text{p-value}_{\mu\neq 10}} = p(|t| > |T|\,|\,H_0) = \left(\int_T^{\infty} + \int_{-\infty}^{-T}\right) \frac{\Gamma(8)\left(1+\frac{t^2}{15}\right)^{-8}}{\sqrt{15\pi}\;\Gamma(\frac{15}{2})}\,dt \approx \boxed{0{,}155}$$

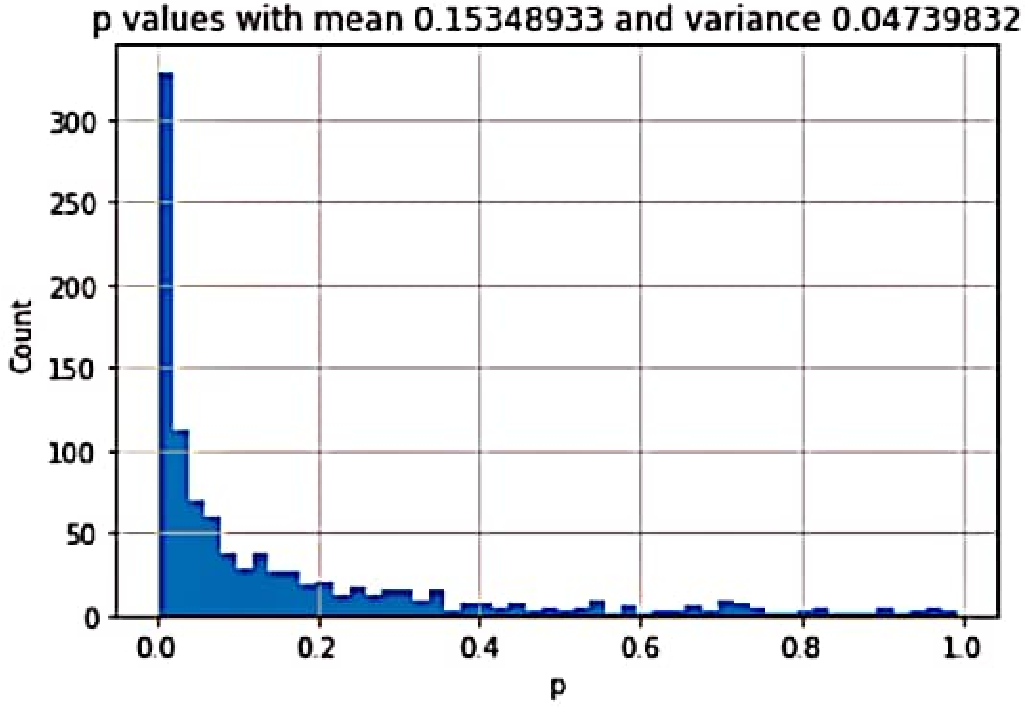
b)
$$\boxed{\text{p-value}_{\mu>10}} = p(t > T\,|\,H_0) = \int_T^{\infty} \frac{\Gamma(8)\left(1+\frac{t^2}{15}\right)^{-8}}{\sqrt{15\pi}\;\Gamma(\frac{15}{2})}\,dt \approx \boxed{0{,}087}$$

<u>We do not reject H_0</u>, same as in 3.1)

```python
#Plotting p_values distribution
h = plt.hist(p_values,50)
#Computing expected p value
counter=0
p_mean= np.mean(p_values)
p_var=np.var(p_values)

#PLOT

plt.grid(True)
plt.xlabel('p')
plt.ylabel('Count')
plt.title("p values with mean {:.8f} and variance {:.8f}".format(p_mean,p_var))
plt.show()
```



We have obtained that p-value $> \alpha$, therefore, we do not reject the null hypothesis.

In[95]:=
```
b = 0;
For[i = 0, i < 500, i++,
  data = RandomVariate[NormalDistribution[11, 2], 16];
  x = (Mean[data] - 10) / (StandardDeviation[data] / Sqrt[16]);
  a = (Integrate[(1 + t^2/15)^-8, {t, x, ∞}] + Integrate[(1 + t^2/15)^-8, {t, -∞, -x}])
    Gamma[8] / (Sqrt[15 π] Gamma[15/2]);
  b = b + a;
 ];
Print[b/500]
```

$$b = 0;$$

$$a = \left(\int_x^{\infty}\left(1+\frac{t^2}{15}\right)^{-8} dt + \int_{-\infty}^{-x}\left(1+\frac{t^2}{15}\right)^{-8} dt\right)\frac{\text{Gamma}[8]}{\sqrt{15\pi}\ \text{Gamma}[15/2]};$$

```
0.155319
```

4)

a)

When the 3 radar guns measure the speed of the car, we will have theree results that come from a Gaussian (μ = speed of car, $\sigma = 5$).

- We will have the Null hypotesis H_0, be $\mu_0 = 40$ m/h
- And the alternative will be then $\mu > 40$ m/h (speeding)

I supose we will be asked to compute p-values given the three measures with unknown μ:

"How likely was that the car was speeding, given this the mesurements"
$\parallel$
NHST: Compute 1-pvalue of a test for the μ value given 3 data points

The null hypotesis is simple, μ is totally determined $\mu = \mu_0 = 40$, the alternative in the other hand is composite, it says only "it's speeding - $\mu > 40$".

(The other trivial alternative hipotesis would be, it is not speeding, but it's exactly the conjugate of the first one, so it would be useless to compute.)

b)

i)

We are going to use a z-test for the μ values as we said in a):

$$z = \frac{\hat{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\hat{x} - 40}{5/\sqrt{3}}$$ which will follow a Null distribution $N(0,1)$

If we want to constrain our "false negative" tickets to a 4% we will need that the $p(\text{type I error}) = \alpha \leq 0{,}04 \longrightarrow$ set $\alpha = 0{,}04$, so:

$$p(z < -z \mid H_0) = \int_{-\infty}^{-z} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz = \frac{1}{2}\left(1 + Erf\left(\frac{-z}{\sqrt{2}}\right)\right) < 0{,}04 \longrightarrow z > 1{,}75$$

And if z fullfills that, then $\hat{x}$ fullfills:

$$\frac{\hat{x} - 40}{5/\sqrt{3}} > 1{,}75 \longrightarrow \boxed{\hat{x} > 45{,}054 \text{ m/h}}$$

ii)

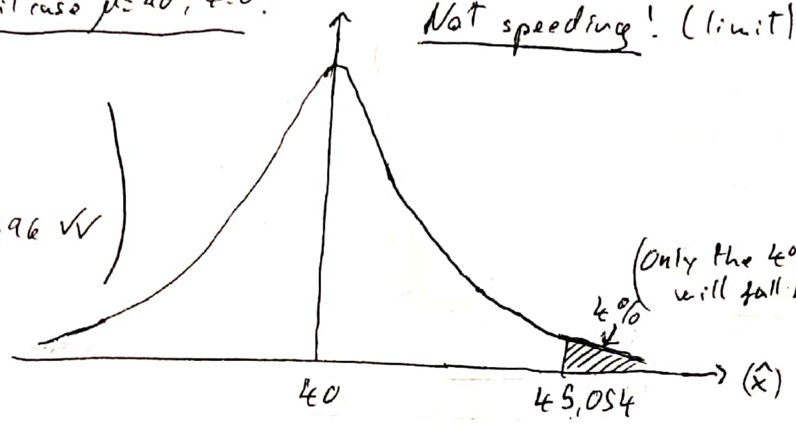
We are using the limit case $\mu = \mu_0 = 40$ because all the rest that have μ (speed) ≤ 40 will recieve even loss "fase negative", so we only need to do the limit case, so:

Limit case $\mu = 40$, $\hat{z} = 0$:

Not speeding! (limit)

(Check:

$$\int_{-\infty}^{45,054} \frac{e^{-\frac{(x-40)^2}{2 \cdot \frac{25}{3}}}}{\frac{5}{\sqrt{3}}\sqrt{2\pi}}\, dx = 0,96 \ ✓✓)$$



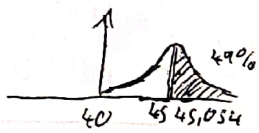
c)

If H_A: $\mu = 45$ m/h, the power will be given by:

- $$\boxed{power_{H_A}} = p(\text{rejecting } H_0 \mid H_A) = p(z < -1,475 \mid H_A) =$$

$$= \int_{-\infty}^{-1,75} \frac{e^{-\frac{(z+\sqrt{3})^2}{2}}}{\sqrt{2\pi}}\, dz = \boxed{0,493} \qquad \left(\frac{+5}{5/\sqrt{3}} = +\sqrt{3}\right)$$

We would give tickets to 49% of the cars speeding at 45 m/h.



To obtain a power of 0,9 with $\alpha = 0,04$, we need to change n until:

- $$power_{H_A} = \int_{-\infty}^{-1,75} \frac{e^{-\frac{(z+\sqrt{n})^2}{2}}}{\sqrt{2\pi}}\, dz > 0,9 \qquad \left(\frac{45}{5/\sqrt{n}} = \sqrt{n}\right)$$

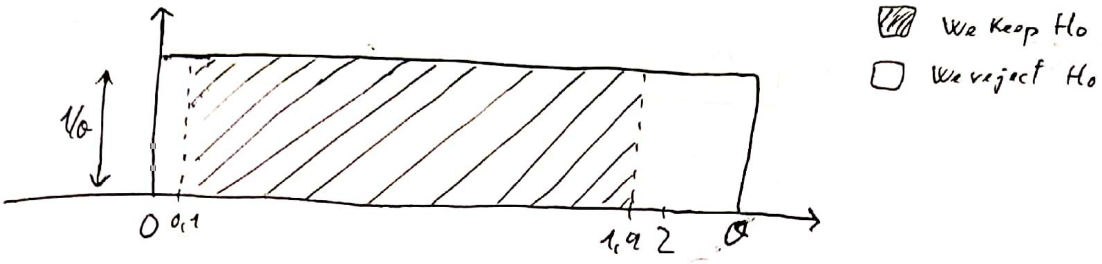
$\longrightarrow$ $\boxed{\begin{array}{l} n \approx 9 \text{ with power } 0,894 \\ n \geq 10 \text{ for power } > 0,92 \end{array}}$

With such number of camaras the threshold on $\hat{x}$ would be:

- $$\frac{\hat{x} - 40}{5/\sqrt{9}} > 1,75 \longrightarrow \boxed{\hat{x} > 42,92 \text{ m/h}}$$

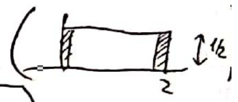
5)

$X \in [0, \theta]$ uniformly with $\begin{cases} H_0: \theta = 2 \\ H_A: \theta \neq 2 \end{cases}$, where we reject H_0 if $X \leq 0{,}1$ or $X \geq 1{,}9$:



a)

$\boxed{p(\text{type I error})} = p(\text{false negative}) = \alpha = p(\text{reject } H_0 \mid H_0) =$



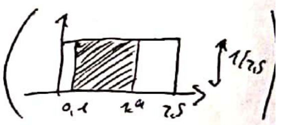
$= p(X \leq 0{,}1 \text{ or } X \geq 1{,}9 \mid \theta = 2) = 0{,}2 \cdot \frac{1}{2} = 0{,}1 = 10\%$

b)

$\boxed{p(\text{type II error} \mid \theta = 2{,}5)} = p(\text{false true} \mid \theta = 2{,}5) = 1 - \text{power}_{\theta = 2{,}5} =$

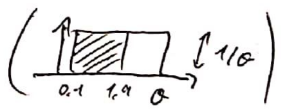
$= p(\text{no reject } H_0 \mid H_A \text{ and } \theta = 2{,}5) = 1 - p(\text{reject } H_0 \mid \theta = 2{,}5) =$

$= (1{,}9 - 0{,}1) \cdot \frac{1}{2{,}5} = \frac{1{,}8}{2{,}5} = \boxed{0{,}72 \approx 72\%}$



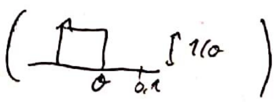
for any θ:

if $\theta \geq 1{,}9$ $\quad p(\text{type II error}) = \frac{1{,}8}{\theta}$



if $0{,}1 \leq \theta \leq 1{,}9$ $\quad p(\text{type II error}) = \frac{\theta - 0{,}1}{\theta}$

if $\theta \leq 0{,}1$ $\quad p(\text{type II error}) = 0$



6)

We have a data with $N = 12$, which has $\bar{x} = \frac{1}{12}\sum_{i=1}^{12} x_i \simeq 0{,}0500...$, from this we can compute the data variance $s^2 = \frac{1}{12-1}\sum_{i=1}^{12}(x_i - \bar{x})^2 \simeq 2{,}34$.

$(\sigma = 1{,}5296...)$

So now let's apply the chi-square test of the link with significance level 0,05:

$$T = (N-1) \cdot \frac{s^2}{\sigma_0^2}$$ where σ_0 is the std. of our null hypotesis H_0.

In this case we are considering $\sigma_0 = 1$, so $T = 11 \cdot s^2 = 25{,}737$

Because the alternative hypotesis is an upper one-tailed one $(\sigma > \sigma_0)$, in order to reject H_0 in favor of H_A, we need $T > \chi^2_{1-\alpha, N-1} = \chi^2_{0.95, 11} = 19{,}675$.

(from table of link)

Answer:

The stadistic T is bigger than the critical value χ^2, so we do reject $\underset{(\sigma=\sigma_0)}{H_0}$ in favor of H_A $(\sigma > \sigma_0)$.

---

7)

Not all distributions will follow Benford's law, depends on what experiments Ada and Carles are doing:

- Rolling dices and summing their values don't need to follow B.L.
- Rolling dices and multiply their values will most likely follow B.L.

But if we assume our phisicist friend, is right, and their experiments should follow B.L., then we can perform a chi-square test, to check this:

- I'll use python to do this using scipy.stats.chisquare(data1, data2) where data1 will be the Ada and Carles provided data, and data 2, will be the Benford's law perfect data example (30,1%, 17,6%, 12,5, ..., 4,6%):

- $\boxed{\text{p-value} = 9.36 \cdot 10^{-13}}$ which is $\lll 0{,}001$ We can be sure their data doesn't follow Benford's law at all.

If the experiments they are conducting had to follow B.L. I would tell the math department Ada and Carles are commiting fraud.