

Hardware Security
Prof. Debdeep Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture – 29
Power Analysis- V

So, welcome to this class on Hardware Security. So, we shall be continuing and finishing off our discussion on how to guess or estimate the number of power traces to estimate the statistical parameters which are required for a difference of mean attack.

(Refer Slide Time: 00:23)



(Refer Slide Time: 00:29)

Estimating the Number of Traces to Estimate μ

- Suppose, we need to estimate the mean with precision c (say 0.01).
- $P(|\bar{X} - \mu_0| > c) = \alpha$, $H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$
- Or, $P(|\bar{X} - \mu_0| > c) = P\left(|\bar{X} - \mu_0| \frac{\sqrt{n}}{\sigma} > c \frac{\sqrt{n}}{\sigma}\right) = P(|Z| > c \frac{\sqrt{n}}{\sigma}) = 2P\left(Z > c \frac{\sqrt{n}}{\sigma}\right) = \alpha$
- Or, $P\left(Z > c \frac{\sqrt{n}}{\sigma}\right) = \frac{\alpha}{2} \Rightarrow c \frac{\sqrt{n}}{\sigma} = Z_{1-\alpha/2}$
- This implies, $n = \frac{\sigma^2}{c^2} Z_{1-\alpha/2}^2$

The slide also features a hand-drawn normal distribution curve with a vertical line at $c \frac{\sqrt{n}}{\sigma}$ and the area to the right shaded, labeled $P(Z > c \frac{\sqrt{n}}{\sigma}) = \frac{\alpha}{2}$. The Swamyam logo is visible at the bottom left.

So, we will be basically like we were in the last class we were discussing about this equation. So, we were basically trying to estimate the; the number of observations which are required to estimate μ properly.

So, if you remember right what we are saying is that, we have essentially you know like a standard normal distribution in which we are trying to. We are basically through which we are trying to use its quantiles to kind of estimate the μ which is there in our normal in our normal distribution. So, we are basically transforming that into a Z statistic and from there we are trying to estimate that.

The idea was that if you have got if you want to you know like estimate the mean with precision c . So, this c could be say 0.01 or 0.001 that depends upon the amount of accuracy that you would like. So, if I want therefore, write I mean essentially α as we discussed in the last class our error probability and that means, that we have got a standard normal distribution and let me just write or draw it once more; so you have got this as your standard normal distribution.

And you are basically trying to estimate the μ right. So, the idea is that if your Z statistic is you know like given more than a specific value; that means, if you are basically trying if your estimate of μ is given as \bar{X} because you do not know what is the actual value of μ and you find out the error.

So, this is your error right because you have you are trying to estimate with that whether it mean the mean is equal to μ_0 . And so therefore, the null hypothesis that you are considering is $\mu = \mu_0$ and the alternative hypothesis is $\mu \neq \mu_0$, but you do not know μ_0 right.

So, you are basically trying to estimate. So, you actually know μ_0 , but you do not know what is the value of μ . So, you are basically trying to estimate it by your by this parameter \bar{X} . So, the idea is that if this particular value that is $\bar{X} - \mu_0$ the absolute value of this is greater than c right then that is an error ok. So, maybe you are basically trying to find out the probability that. So, you know that this probability is equal to α .

So, now from here we basically try to derive this equation which we derived in the last class and we got this as a corresponding estimate that is n is given by $\frac{\sigma^2}{c^2} \times Z_{1-\frac{\alpha}{2}}$ whole square. But often even we analyze power attacks we do not you know like need to actually estimate the value of the mean, but what we rather require is to estimate whether the means are different or whether the means are same ok or we need to kind of estimate whether the mean is say less than 0 or the mean is greater than 0 ok.

So, we basically need to estimate the mean is say equal to 0 or mean the mean is not equal to 0. So, therefore, right we probably need to do an this is another kind of problem right which we need to also kind of solve.

(Refer Slide Time: 03:35)

$$\begin{aligned}
 &P(\bar{X} < 0) = 1 - \alpha, \text{ given } \mu = \mu_0, \text{ with } \mu_0 < 0 \\
 &\downarrow \\
 &P\left(\frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} < -\mu \frac{\sqrt{n}}{\sigma}\right) = 1 - \alpha \\
 &\quad \quad \quad \downarrow \\
 &\quad \quad \quad Z \sim N(0, 1) \\
 &\Rightarrow P\left(Z < -\mu \frac{\sqrt{n}}{\sigma}\right) = \phi\left(-\mu \frac{\sqrt{n}}{\sigma}\right) \quad \left| \begin{array}{l} \phi(z_\alpha) = \alpha \\ \phi(z_{1-\alpha}) = 1 - \alpha \end{array} \right. \\
 &\quad \quad \quad \downarrow \\
 &\quad \quad \quad -\mu \frac{\sqrt{n}}{\sigma} = z_{1-\alpha} \Rightarrow \mu^2 \frac{n}{\sigma^2} = z_{1-\alpha}^2 \\
 &\quad \quad \quad \downarrow \\
 &\quad \quad \quad n = \frac{\sigma^2}{\mu^2} \times z_{1-\alpha}^2
 \end{aligned}$$

So, we basically are trying to kind of analyze this fact that P of say X bar which is your estimated mean; whether this is lesser than equal to 0. And we know that we are as we know that this probability right is equal to 1 minus alpha. So, we are basically trying to say that with a good amount of confidence and the confidence is 1 minus alpha; we are trying to find out the probability that X bar is lesser than 0; suppose your mean is negative. So, this is given that the mean or mu equal to mu naught with mu naught lesser than 0.

So, in this case in this case the mean is a negative value. So, that means, right if you if you observe the corresponding. So, we can basically rewrite this in equation by you know like just writing X bar minus mu into root n by sigma and writing that this is lesser than equal to minus mu root n by sigma ok; that is basically just I am writing X bar I am deducting a minus mu from both sides; so, I get minus mu and then I multiply by a root n by sigma.

Again note that sigma is positive; so therefore, if I multiply then a signs do not change. So, this probability right essentially remains the same and essentially is also equal to 1 minus alpha. So; that means, right this part essentially is nothing, but my Z statistic which essentially follows a standard normal distribution. So, essentially the mean is 0 and sigma is 1 ok; this is what we have already seen in the last class. So; that means, right this we can write it as P Z is lesser than minus mu into root of n by sigma.

And that is nothing, but the cumulative frequency at the point or few you know Φ evaluated at $-\mu/\sigma\sqrt{n}$. And; that means, right if you remember right we said that so, this is equal to $1 - \alpha$. And we discussed right therefore, that this parameter that is $-\mu/\sigma\sqrt{n}$ is nothing, but $Z_{1-\alpha}$ right because as we define right that $\Phi(Z_{\alpha}) = \alpha$ ok.

So; that means, $\Phi(Z_{1-\alpha}) = 1 - \alpha$ right. So, therefore, we have got this and note that your mean or μ right is essentially negative and therefore, what you essentially have is this equation. So, now from here I can estimate by squaring; so if I square, I will get μ^2/n divided by σ^2 and that is equal to $Z_{1-\alpha}^2$ or we basically get n is equal to σ^2/μ^2 by $Z_{1-\alpha}^2$.

So, this is a different equation than what we got because our objective is different. And very often you may find that this n is probably much lesser than the previous value of n ok. And you can try to you know like use some possible values and from there you can try to experiment that whether that ordering is true or not; this I leave it to you as an exercise.

So, so therefore, right I mean with this background right we can now get back again to our discussion and continue ok. So, therefore, right this is what I just wanted to show is that this number of trace right depends upon the hypothesis. And it is not that whatever hypothesis you make this is the number of observations required ok; it will depend upon your hypothesis. So, you need to do a case by case analysis ok; so it is interesting how you derive them and depending upon the hypothesis which you make.

(Refer Slide Time: 08:05)

Estimating the number of traces for estimating DoM

- Consider: $X \sim \mathcal{N}(\mu_X, \sigma)$, $Y \sim \mathcal{N}(\mu_Y, \sigma)$, with different means but same variances. $\sigma^2 = \frac{\min(m,n)}{mn}$ $\sqrt{\frac{2n}{n^2}} = \sqrt{\frac{2}{n}}$
- Then, $\bar{X} - \bar{Y} \sim \mathcal{N}(\mu_X - \mu_Y, \sigma \sqrt{\frac{n+n}{n^2}})$, assuming there are n traces in each group.
- Then, $Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sigma \sqrt{\frac{2}{n}}}$ is a standard normal distribution.

So, so therefore, right we are now almost said to estimate the number of traces which I need for estimating the difference of mean. So, now we have got to you know like distributions ok.

So, we have got a distribution 0 and we have got a distribution 1 right. So, we have got in our difference of mean attack we are essentially having 2 bins. So, we have a 0 bin and we have a 1 bin right. So, what we do is we are basically having our encryption algorithm say let me call this as an AES round for example.

And in the final operation right you basically do an exclusive or with your tenth round key for example; if you are talking about AES 128 and this is the corresponding cipher which you are generating. So, what you are doing as an attacker is that from this key you are basically guessing only a portion of this key right. And then you are basically targeting say the last round say one of the S boxes here.

So, you are basically like inverting this S box going to a specific bit and if this bit right depending upon your guess here that is the guess that you make for this key bite is you know like whether this according to this guess you get a 0 value here or you get a 1 value here from the cipher text. So, you basically take the cipher text you XOR with the key you take the inverse S box and you go to that target bit right and depending upon whether the target bit is 0 or 1 you put the power trace into the 0 bin or you put the power trace into the 1 bin.

So, now we have got 2 statistical distributions; one distribution say let me call it X and the other distribution is what we have called as Y. So, X actually follows a normal distribution; so this normal distribution is essentially say μ_X comma σ and Y essentially follows also another normal distribution say denoted as μ_Y and σ ok.

So, note that I have kept the sigma same in both the cases which is most likely to happen ok. Although there may be small differences in that case I will probably substitute by an average sigma because they are roughly the same because they are taken from the same population and likely they would be same. So, I am just definitely assuming that they are same for simplicity of our analysis.

So, therefore, right now the interesting thing is that what we need is to estimate this distribution or \bar{X} minus \bar{Y} . That is you see that I have taken the sampling distribution of the average of this distribution; I have taken the sampling distribution of the average of this distribution and then I am considering the difference of them ok.

So, that is denoted as \bar{X} minus \bar{Y} this will also follow a normal distribution where the mean will get reduced; the mean will get subtracted μ_X minus μ_Y , but the interesting thing is that the variance right will get added up. So, for example, like if the sigma in this case I have assumed that both the traces; I have got same number of values same number of samples, if it is not the same then you can probably throw away some samples and make it same ok.

But if you are also if you, but you can also generalize it you say that this has got say m traces and this has got say n traces; then this sigma would be sigma square root of m plus n divided by m n. So, if I substitute m and n m and n as same then basically you have got sigma into square root of m plus n divided by 2 n and this right means that it is essentially I mean it basically; so it is basically m plus n divided by n squared. So, maybe you should correct it actually this is m plus n by n square.

So; that means, right it is square root of 2 n divided by n square and that is equal to nothing, but square root of 2 by n. And that is essentially this denominator factor which is sigma square root of 2 by n. And the now if I make a transformation that is I basically take \bar{X} minus \bar{Y} and I subtract out the mean of this distribution which is minus μ_X minus μ_Y and divided by sigma square root of 2 by n; which is the standard

deviation of this distribution then as expected right like what we discussed before we get a standard normal distribution.

So, now we will try to play around with this Z and again develop the confidence intervals for mu X minus mu Y ok; so because that is what we want to estimate. So, therefore, right I mean so with this background let us see how we can do that.

(Refer Slide Time: 12:45)

Confidence Interval for DoM

- Confidence Interval for $\mu_X - \mu_Y$ for known σ is by rewriting, $P(Z_{\alpha/2} \leq Z \leq Z_{1-\alpha/2})$:

$$\left[\bar{X} - \bar{Y} - \frac{\sigma}{\sqrt{n/2}} Z_{1-\alpha/2}, \bar{X} - \bar{Y} + \frac{\sigma}{\sqrt{n/2}} Z_{1-\alpha/2} \right]$$
- Null Hypothesis, $H_0: \mu_X - \mu_Y = 0$, and Alternative Hypothesis, $H_1: \mu_X - \mu_Y \neq 0$.
- We test if $Z = \frac{\bar{X} - \bar{Y}}{\sigma \sqrt{\frac{2}{n}}}$ is in the critical region.
- Thus, as $2P \left[\frac{\bar{X} - \bar{Y}}{\sigma \sqrt{\frac{2}{n}}} > \frac{c}{\sigma \sqrt{\frac{2}{n}}} \right] = \alpha \Rightarrow \frac{c}{\sigma \sqrt{\frac{2}{n}}} = Z_{1-\alpha/2} \Rightarrow n = \frac{2\sigma^2}{c^2} Z_{1-\alpha/2}^2$

So, we basically try to or want to derive the confidence interval for difference of mean that is our objective. The confidence interval for mu X minus mu Y for a known sigma therefore, you can derive from this confidence interval for Z right. So, the confidence interval for Z as we discussed in the previous class was from Z alpha by 2 to Z 1 minus alpha by 2. And now we make again this substitution that Z essentially is nothing, but this that is X bar minus Y bar minus mu X minus mu Y divided by sigma square root of 2 by n.

Now note that in this; so basically like if we make a hypothesis that you know like. So, basically like if you make so what we do is that I mean if the corresponding confidence interval for mu X minus mu Y; in that case side will be X bar minus Y bar minus or you can say X bar minus Y bar plus minus sigma by square root of this would be 2 by n actually ok.

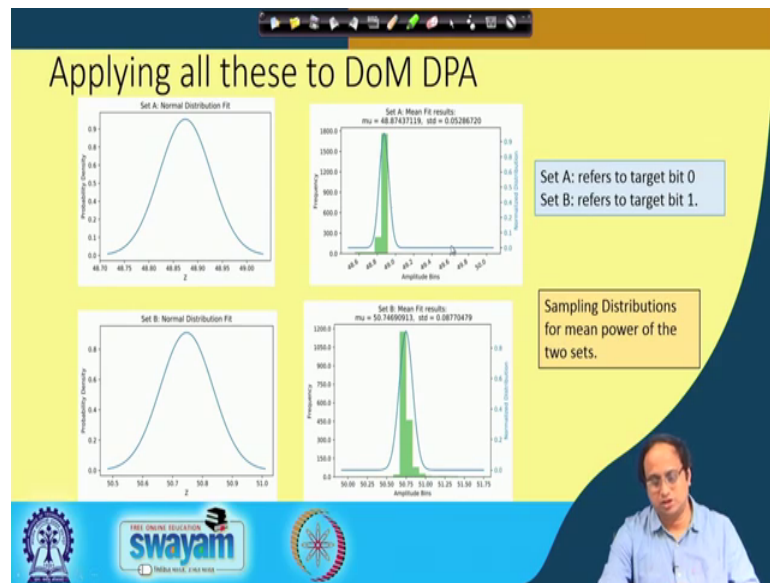
So, maybe again you should make another connection which is that this is square root of 2 by n . So, square root of 2 by n ; so plus minus sigma by square root of 2 by n Z $1 - \alpha$ by 2 . So, now we make a null hypothesis the null hypothesis is essentially that $\mu_X - \mu_Y$ is equal to 0 which means that both the means are same the alternative hypothesis is that $\mu_X - \mu_Y$ is not equal to 0 .

So, therefore, right we test this statistic Z . So, note that here I have assumed the null hypothesis; so therefore, $\mu_X - \mu_Y$ is 0 and therefore, I test that if Z is equal to $\bar{X} - \bar{Y}$ divided by sigma square root of 2 by n whether this lies in the critical region. So, the critical region is essentially as previously depicted. So, you can again apply the same technique.

So, what you basically have got is 2 times the probability that $\bar{X} - \bar{Y}$ by sigma square root of 2 by n that is this value is basically greater than c by sigma square root of 2 sigma square root of 2 by n . So; that means, that $\bar{X} - \bar{Y}$ is exceeding c and that is c is essentially the precision or the tolerance which you are allowing.

So, therefore, right if I said this as α because this is essentially α in this case. Now; that means, that this parameter that is seek c by sigma square root of 2 by n is nothing, but Z $1 - \alpha$ by 2 ; as we have previously discussed. And therefore, now the number of traces can be gone obtained by this formula. So, n is equal to 2 times sigma square by c square into Z $1 - \alpha$ by 2 whole square. So, therefore, write the again right with this equation; so now, we will be trying to you know like you utilize this in an actual experimental setting.

(Refer Slide Time: 15:45)



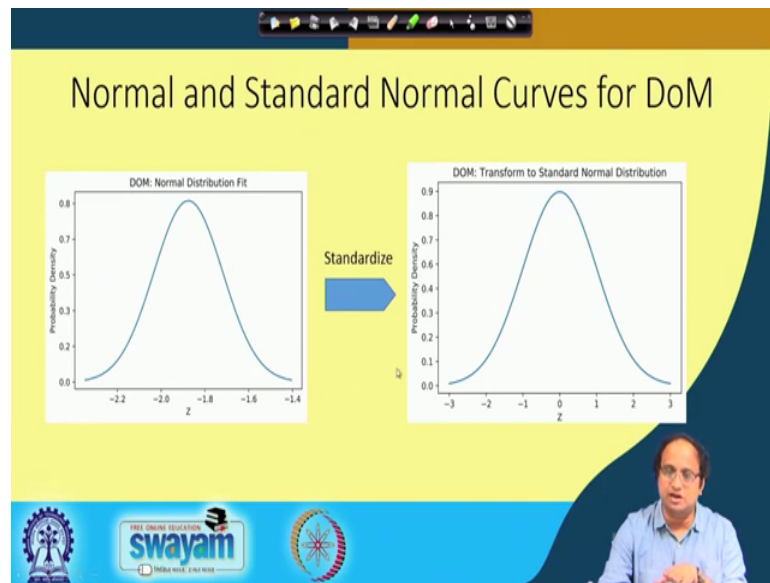
So, therefore, right let us see this case study. So, this case study shows how you are; so we have is again you know like trying to perform a difference of mean at that on the last round of AES.

So, in the last round of AES again we have got 2 distributions because of the 0 bin and the 1 bin that I talked about. So, suppose this is your normal distribution; I mean the fitted normal distribution for one of the sets I call that a set A for example, and this is the sampling distribution for mean ok. So, note that the variance essentially has got reduced and that is essentially expected right because the variance should get like sigma by square root of n.

So, therefore, we need we need get a narrow normal distribution and likewise right for the other being you have you have this normal distribution. And if we again apply or obtain the sampling distribution for the mean you again obtain another set which is again narrow ok. So, therefore, the set A refers to target bit 0 and set B refers to target B bin 1 ok; the target to get B this is the same target bit, but sometimes it takes the value of 0 and sometimes it takes the value of 1.

So, now we observe that these are my sampling distributions that so; that means, that this is your \bar{X} set and this is your \bar{Y} set ok. These are the sampling distributions for the mean which belongs to set A and this is the sampling distribution for the set B or the mean of set B.

(Refer Slide Time: 17:06)

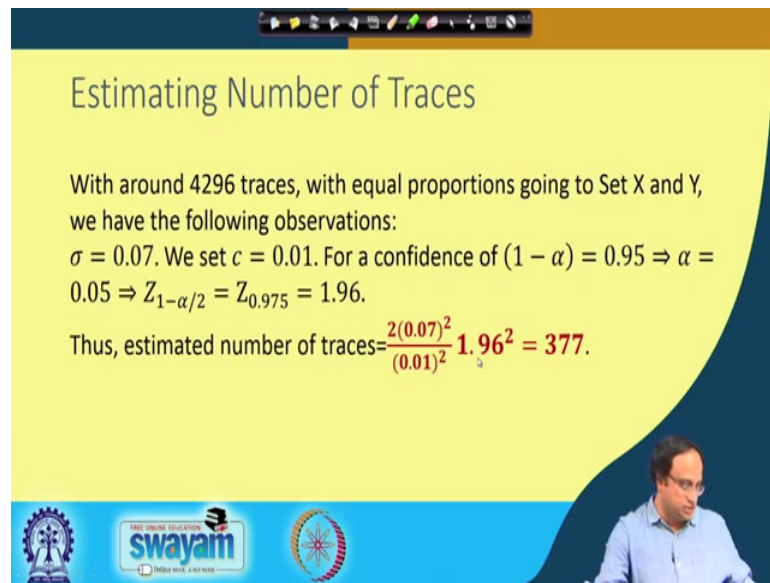


So, now what we just discuss here that you essentially you take one of these difference of mean plots. So, this is the difference of mean plot.

So, after the difference of mean now takes a nice normal distribution. So, the difference of distribution means that you have got all these averages and you are basically trying to take the difference of these averages ok. So, you are taking the difference you are taking an average from one set you are taking the average from the other set and you are computing their differences.

So, this difference right as we discussed that $\bar{X} - \bar{Y}$ also follows a normal distribution and this is how it looks like in our case ok; so this a normal distribution. You if you want to standardize it again you can apply the standardization technique and this is the corresponding standard normal distribution curve.

(Refer Slide Time: 17:53)



Estimating Number of Traces

With around 4296 traces, with equal proportions going to Set X and Y, we have the following observations:
 $\sigma = 0.07$. We set $c = 0.01$. For a confidence of $(1 - \alpha) = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow Z_{1-\alpha/2} = Z_{0.975} = 1.96$.

Thus, estimated number of traces = $\frac{2(0.07)^2}{(0.01)^2} 1.96^2 = 377$.

The slide also features logos for Swamyam and other educational institutions, and a small inset image of a presenter in the bottom right corner.

So, now the final thing we want to estimate the number of traces which you are required. So, what we do is that, we observe here around 4296 traces with roughly equal proportions going to set X and set Y.

For our analysis we have discarded some cases where you know like there are aberrations and things like that we have discarded few cases where we had inequality in these two sets. So, that finally, we have got two sets with roughly the same number of any number of samples. And a sigma it turns out to be 0.07; again the sigma was not exactly 7. So, we kind of kind of approximated by the average they were almost same. So, we approximated by the average sigma and we set a c value of 0.01 that is the accuracy that I allow ok.

That means μX minus μY is equal to 0 is my null hypothesis I allow a small tolerance of say 0.01; that means, it can be 0.01 above 0 about below 0. So, therefore, right for a confidence of suppose I set a confidence of 0.95; suppose somebody asks you that you should guess this with the confidence of 0.95; that means, 1 minus alpha is 0.95 right that implies that alpha is equal to 0.05 and; that means, 1 minus alpha by 2 is equal to 0.975 right.

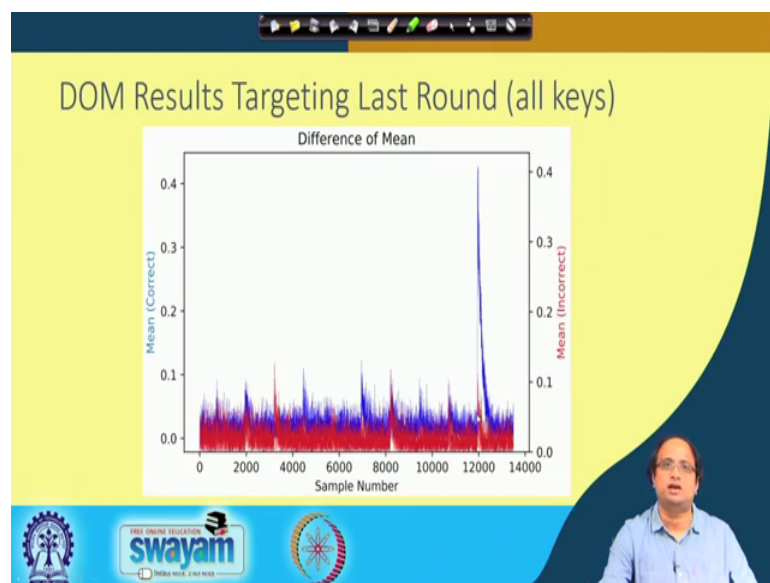
So, therefore, you should find out the quantile at Z 0.975 and if you refer back to the table that we saw the corresponding entry over there was 1.96. So, therefore, right this is my corresponding value here 1.96. So, therefore, if I just plug in here to my equation

which is essentially given in the previous you know discussion that is n is equal to $2 \sigma^2 / c^2$ into $Z_{1-\alpha/2}$ whole square.

So, we just plug, plug in these values here to get 2 into 0.07 square by 0.01 squared into 1.96 square and this works out to 377 ok. So, again this is a rough approximation, but you know like you get a rough estimate that the number of traces which are required to indeed understand or accept the null hypothesis with a reasonably high confidence.

So, if we are accepting the null hypothesis right then what does it mean? It means that $\mu_X - \mu_Y$ is roughly equal to 0 . So, in which case will it convey it will correspond to the wrong guess or the current guess; it will correspond to the wrong guess. Because in the wrong guess both the bins right will have the same average or similar averages, but if you are doing it for the correct guess right where you are basically kind of rejecting the null hypothesis those are your candidates case ok. So, that is exactly what we will do in our attack.

(Refer Slide Time: 20:30)



So, what we do in our attack is basically we carry on the attack eventually and you can observe that as I say that we are trying to target in the last round of AES. So, these are all the rounds that we have targeted. So, this is the corresponding you know like the samples that we have observed.

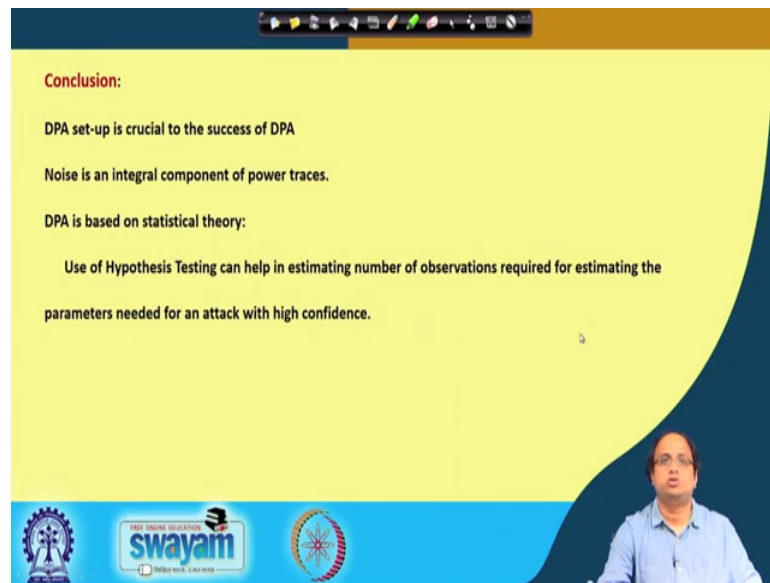
And the red are the red colour has been used to indicate the wrong guesses and the blue for the correct one. And you can see that the difference of mean for the correct right essentially is kind of strikingly you know distinct from the wrong ones ok. And you essentially are able to distinguish the correct one the correct key guess from the wrong key guesses ok. So, this roughly corresponds to the last part of your cipher where you are doing the last round of your encryption.

(Refer Slide Time: 12:12)



So, this is the reference that we have followed. So, of course, we have followed our standard textbook, but this is also another very nice textbook which you can refer to it is written by Stefan Mangard, Elizabeth Oswald, Thomas Popp it is called Power Analysis Attacks: Revealing the Secrets of Smart Cards this is published by Springer.

(Refer Slide Time: 21:30)



Conclusion:

DPA set-up is crucial to the success of DPA

Noise is an integral component of power traces.

DPA is based on statistical theory:

Use of Hypothesis Testing can help in estimating number of observations required for estimating the parameters needed for an attack with high confidence.

swamyam

So, to conclude what we discussed is DPA setup is crucial to the success of DPA. Noise is an integral component of power traces and therefore, you need to properly analyze noise. The noise can be of two types it can be of electrical noise it can be algorithmic noise as we say discussed the algorithm noise also has got an influence on the algorithm on the architecture; like if you can we will see later on how we can compare a parallel architecture with a serialized architecture with respect to algorithmic noise.

The electrical noise is more fundamental and depends upon the platform ok, but at the same time you would like to reduce the electrical noise by applying proper insulation and proper and using proper probes for community in the power signal from your target to the oscilloscope. And, DPA right is largely based on statistical theory; so it is important to kind of go back and look at these techniques which are like hypothesis testing and so on which can help in estimating the number of observations which are required for you know like guessing properly the parameters which are needed to do an attack with high confidence ok.

So, it is always good to back up our experiments with a good rationale and proper explanations using statistical tools ok. And these statistical tools which we just now discussed helped us in doing so ok.

So, with that I would like to thank you for your attention.