[00:01:00] So we're now recording.
Everybody, welcome to the Monday webinar/consultation session for Week 1.
Also welcome to Week 1 of the course.
The plan today is, I think I will actually hold off on talking about the quiz until Wednesday.
Since then there might be more specific questions.
The quiz should go up later tonight, I believe.
Other than that, I'm just going to go through the questions that folks have posted on the thread and then just work through those.
Any questions before we start? As always, if you have a question, particularly [00:02:00] if you're typing a question, please click the raised hand icon so that I would know the question is coming so I don't move on.
So the first incompletely answered question that comes from Larissa Simpson.
This relates to the two-sample hotel entities T squared test and the case of unequal variances.
I'm going to [00:03:00] share the screen.
One second.
There we go.
So I assume that this is the character you're talking about.
Is that right? This is a bit [00:04:00] of a messy formula and I didn't help because I actually forgot to define a bit of notation in the notes which have now updated.
But here's the thing and I'm not going to derive this formula.
Probably wouldn't be able to, at least off the top of my head.
But I'm going to talk about why do we need this particular formula, this particular quantity.
In your interest as class, you might remember the difference between the Z-test and the t-test.
The Z test, which you use when you know the population variance.
The t-test is when you don't.
The t-test has degrees of freedom which increase as your sample size increases.
If you look at the density of t versus the normal density, [00:05:00] you see the t is a bit wider, which accounts for the fact that when we have to estimate our standard deviation or variance, that adds extra noise to everything we do.
The degrees of freedom over the t-distribution quantifies that noise.
Making them a bit bigger.
So when we talk about the two sample t-squared test, a lot of similar ideas apply because you might remember that for the two-sample t-test, you have a similar complication.
When you assume that, for example, you can pull the variance, that is, you're willing to assume that the population variance covariance matrix in group x is the same as in group y, well then you compute your estimate for the population variance [00:06:00] pooled.
Here we'll take essentially a weighted average of these.
The information you have about the variance is actually the combined information from x is x sample and from y sample.
That means that your effective sample size for estimating the variance is going to be the total sample size between x and y.
N_x plus n_y minus 2 because of the sample means.
In fact, you should look at this expression which were the degrees of freedom of the F distribution we use here, it is based on n_x plus n_y and then try various constants.
Now, the thing is where we use the snooze we were in a non-pooling setting.
If we're not prepared to make the pooling assumption.
Here, instead of having one estimate as capitalist [00:07:00] pooled, we have a separate variance estimate for x and for y or variance-covariance matrix estimate.

We need to figure out how much additional noise comes from the fact that this bit in the middle is in fact estimated rather than known exactly.
Again, in the pooled case, it's straightforward.
In the unpooled case, it turns out to be complicated.
But the point is, find a place where I can select it here, well, instead of this n_y plus x plus n_y minus 2, we have this blue.
This is basically a measurement of how much information do we have about [00:08:00] the variance.
I guess the smallest point is the more additional noise we have due to having to estimate the variance.
Now, as it turns out, this is the formula you end up with.
It's a function of a whole bunch of things.
Again, it's probably not worth deriving, but I think it just was an intuition.
Well, first of all, the first thing we can see here, if we look at this formula, well, you have the dimensions of the problem, [00:09:00] we have the sum from i equals 1- 2.
Basically that's a shorthand for all the both S_X and S_Y because we want to repeat this expression twice.
The first thing we hear is one over n_i.
Now, that means that because this is something divided by something divided by something, the bigger the n_i is going to be, the bigger the numerator is going to be.
Does that make sense? Can I get a quick show of hands, who is familiar with the idea of harmonic mean? [00:10:00] May be I'll switch to the document camera, just briefly write it down here.
Hopefully, it's right-side up this time.
The idea of a hallmark mean, it's one of the many different ways to summarise the central tendency of a bunch of numbers.
The idea there is that you write it as the, let's say x^h [00:11:00] is going to be n over the sum I equals 1 to n of x or i inverse, i or 1 over x_ i.
Let's write it like that, 1 over x_i.
The idea is that the bigger each of the x size is, the bigger the harmonic mean is going to be, but also it's not, and in fact, the final result will be somewhere between the minimal x_i and the maximum x_i.
But [00:12:00] it turns out to be weighted towards the smaller of the values.
Just imagine, for example if one of these x_i's is really tiny, then 1 over that x_i is going to be really huge.
The denominator is going to be huge, and so the harmonic mean is going to be small.
This is actually what we're seeing here.
Let me switch back to sharing screen.
No.
Okay.
I think it's this one.
Okay.
Yeah.
Yes, good.
It's sharing.
What we see here is that in fact, what most [00:13:00] of your determines is new is the smaller of the n's.
Because if one of the n's is very small, then there is only so big that this u can be.
Otherwise down here it ends up, I think, being a function of how different these two variances are.
In fact there, if the s_i's are similar, then it's going to increase u by a little bit.
And again, I don't think I can give you any better intuition about this but that's what we have.
Now how do we implement this? I guess the best answer is very carefully, because ultimately these are

just simple mathematical operations.
I [00:14:00] wish I could actually share both the screen and the document camera at the same time, but I guess that's not something we can do here.
But here, for example, you would probably just have the first sample variance over n, one second sample variance over n two.
Then add them up and then use solve function.
I'll type them into chat.
Then that should do it.
Use their sole function to take the inverse of that resulting matrix, then multiply it by s_i using the matrix product.
Then square it just by multiplying the matrix by itself.
Now when you take a trace of the matrix, there is no trace function in R, but [00:15:00] actually it's pretty simple.
The trace is just the sum of the diagonal of M, where M is the matrix you want.
Then over here, again, it's similar, but you switch the order here.
It's the trace of the square, and here, it's the square of trace.
But then you do that for each of the two matrices.
Does that make sense? I know, this is not a very nice sensor, but again, I don't think I'd be able to drive it not without going very deeply into theory and sorting it out.
But if you'd like, I can give you a reference [00:16:00] for this.
All right.
Did that help? Yeah.
I think one thing is that I think I will do is I'll post some hints about the challenge with some functions you might want to look up that might make your job easier as well.
That was first question.
Now, I'll look question list here.
I have two monitors and it's still not enough display space for everything.
[00:17:00] All right.
The next question and let's see this one was about something.
There was a question from Shane Lavington about when I'll be releasing extra solutions each week? Yes, I will.
Again, I prefer to release them mid-week to give you a chance to work on those.
But I meant to answer that on the board anyway.
Sheevia.
All right.
Next question is also for Larissa Simpson to go through the task 1 of the metrics exercises.
Let me pull that up and [00:18:00] I'll try to explain because that one did involve a bit of algebra as well as programming.
Now, I do have to say our studio is glitching on my system for some reason and I don't think we've been able to figure out why.
But I will say I'll just pull up my own session and load up the file.
All right, back to sharing screen.
All right.
I think we're okay.
[00:19:00] There we go.
The task one has explained why the empirical covariance matrix with a factor of 1/n minus 1 in front of the 5 variables X1 through X5 can be calculated using the following commands.
That turns out to be a bit of an algebraic question.

First, let's see what this actually does.
What this our statement down here actually does.
Well, first thing is it computes 1/n minus 1.
That's pretty straightforward.
Then we multiply it by this big matrix expression.
Now, the first element is t of x.
Now little t, that's just transpose.
[00:20:00] Then matrix multiplied by something in parentheses here, and then again multiplied by x.
What's in parentheses? First, you have diag n that construct an identity matrix to be precise, that is n by n.
Then the second bit here, let's look at what it does.
We have matrix 1_{nn}, so that's a matrix.
You should be able to see it now.
I'll paste the link into chat.
But you should be able to see it now, it's been up there for a bit for a while now.
Amy can you [00:21:00] confirm that you can access this, because if not, then I need to talk to ed people again.
Good.
Somebody can see it.
That's a good sign.
The next thing is we have a matrix whose elements are 1, recycled as needed.
That is n by n.
We have used n by n matrix of ones.
Then you have a new line then you divide that by n essentially.
I just subtract the identity matrix.
That's a weird layout.
[00:22:00] Let's see.
In fact, let's take a quick look at what that matrix actually looks like.
I'll just use 5 as the size.
On the diagonal you have 1 minus 1/n.
After it you have minus 1/n.
Now, we can interpret this in a number of ways.
[00:23:00] I think I'll just go through this now.
Let's say we have our matrix X, that's the combined matrix of all these vectors.
This is the expression then implemented by r.
Now here I think, let me just try the matrix algebra.
We start with this expression from before, then we can use the distributive property of multiplication.
That is, we bring X transpose into the parenthetical expression here.
Then we bring X into the expression.
We end up with X transpose identity matrix X, we have X transpose matrix of 1, x minus [00:24:00] x/n.
Identity matrix times something is just the same thing, so X transpose X and X transpose 1,1x/n.
Now to see what result we do need to actually look out to the, well, the definition of matrix.
We actually need to write out.
What will be the sij element of this expression? Well, so it's 1/n and then X transpose X, well, just by definition of matrix multiplication, that sum from k minus k equals 1 to n of X transpose ik times xkj.
That's the definition of matrix multiplication, which we can write as since this is x transpose IK, that's the same as [00:25:00] x_{ki}.
No, wait, no [inaudible 00:25:04] Oh, yes.

Sorry.
My apologies.
Yes.
I hear it.
You can.
Sorry.
Well, this is a bit sloppy, although it's still correct.
Because it ends up being symmetric.
[00:26:00] Sorry, my apologies.
Yes.
I think I may have used the notation a little bit here.
The idea is that you take the ik element of matrix X and then transpose it.
But that element is a scalar, so it doesn't matter.
Then you have x_ik which is basically the product of the i's and j's variable k'th observation.
Then here you have the, they hear it, get its little more complicated.
But basically what you end up doing is you will multiply first these two elements.
This one, again, that's just x_ik transpose is just x_ik, because [00:27:00] x_ik is a scalar.
Then here we use the definition of matrix product.
Well, which is the sum over l equals 1 going up to n of, I guess matrix of ones kl element times the element kj.
But, this is, excuse me, I need to fix something here.
But that's just 1.
So we end up with an expression which is the sum of x_ik times x, sum of x_jl, and then divided by n from here.
This, you've probably seen before as a covariance calculation formula.
You sum up, if you prefer, [00:28:00] the sum of xi times yi minus the sum of x's times the sum of the y's divided by n.
That's where this expression comes from.
I hope that helps under Amy Tagore, I'm sorry, I apologise, I didn't see the message.
The link is not working.
Have you logged in to the, okay, good.
There we go, that's basically how this bit works.
Then the next question was from [inaudible 00:28:57] about [00:29:00] 0.
1 and the one about eigenvectors and eigenvalues.
By the way, if I mispronounce your name at any point, please let me know and I will do my best to start pronouncing it correctly.
This theorem is a bit cryptic.
We're not going to prove it.
We're not going to use it.
[00:30:00] In a nutshell, if you have a real symmetric matrix X, these several bits here are important.
If you have a matrix X, which is real, that is supposed to complex or imaginary and symmetric, then it has p eigenvalues, so p by p matrix, and they're all real.
It turns out that if it's not symmetric, sometimes you have situations where eigenvalues are complex.
Now, the next step, the thing is when you have p eigenvalues, but some of them are actually repeated.
Remember the definition of an eigenvalue is this.
Well, it's the own value.
There's a number of ways of saying it, but basically if you have, [00:31:00] it's a solution to this quantity or maybe it's an eigenvalue to value that has this property.

That is, if you have matrix X, if it's symmetric and real then eigenvalues must be real, I believe, yes.
Then eigenvalues, the idea is that when you evaluate this characteristic polynomial, well it's a polynomial.
Well, first-degree polynomial has one solution.
That's basically a line and at some point that line will cross [00:32:00] the x-axis.
The second degree polynomial in Lambda might have two solutions, and so on.
Not all resources there are always real.
That's where that whole issue with if it's not symmetric, then it might not be real.
But anyway, the idea is that because of the way the determinant here is just the way that works, the determinant of a p by p matrix is essentially p degree polynomial in the elements.
When you write down the characteristic polynomial, the degree of the polynomial is going to be p so it will have p solutions.
However, some [00:33:00] of the solutions are duplicated, just like it is possible for a quadratic polynomial to have only one unique solution.
You might recall the quadratic formula.
Negative b plus or minus square root of b squared minus 4ac, everything divided by 2a.
Well, if that bill after plus or minus turns out to cancel out a new zero, well you only have one root.
Makes sense? That's one way of looking at it.
But the other way of looking at it is that really you have minus b plus or minus 0 divided by 2a.
There are really two roots, but they're just the same number.
The same thing happens with eigenvalues.
They can be repeated.
Now, [00:34:00] these are the easy cases when they're not, so that's when they're all different.
Then you can define the eigenvectors uniquely.
These eigenvectors will be orthogonal.
Again, we're going to prove this orthogonality, I think I talked about last week.
Again, these eigenvectors are unique up to a constant.
Because, again, if you take this expression, you multiply the eigenvector by a constant here, the expression still holds.
Now, if the eigenvectors do coincide, they're not necessarily unique.
But they can be chosen to be orthogonal to each other.
An orthogonality again has these properties and I think I did a geometric illustration last week.
[00:35:00] This is basically what this is talking about.
That when you have a real symmetric matrix, all eigenvalues are real.
If they're not, you basically end up with complex roots to the polynomial.
The rest is about eigenvectors and the fact that you can always get these orthogonal eigenvectors.
Does that help or are there any other questions? [00:36:00] Pointed out.
That's just the equation number.
That means that I'll probably refer to it down here.
Yeah, that's just standard notation.
It's just equation numbers that were referred to later.
Next item is the question by Tarika Jackman.
I hope I'm pronouncing that correctly.
Let me pull up the question.
[00:37:00] The question is basically some guidance on the multivariate normal distribution and multivariate distribution challenges, task Number 2.
Finding an appropriate transformation for these variables and check they are now normal.
I think we're going to obviously post solutions later this week but here's some things to think about.

If your distribution is already symmetric then typically you can usually use it for things that require normality, thanks for central limit theorem.
When the distribution is not symmetric, then you can try to transform it [00:38:00] to maybe make it more symmetric.
That's generally the first thing you want to try to do.
In particular, most often the essential we get is something being skewed to the right.
The idea is that you have lots of small values and then a small number of very large values.
A lot of different functions can be used to bring in those large values so that they don't have quite as much impact.
The more modest ones are the square root.
It's a concave-down function like others if that makes any sense.
You have to find its square root and in fact, maybe I'll switch the document camera just to illustrate the point, maybe draw a picture here.
This works.
I'm going [00:39:00] to draw it.
Let's say this is our original distribution.
We have a bunch of observations down here and then one extreme observation here.
Now let's say we have a square root function that is going to be shaped something like this.
What happens when we map this to the square root function? The small values get mapped, relatively speaking, to each other pretty close to where they were.
The big value is still pretty far away but it's been brought in.
It's no longer quite as far away from the others.
Does that make sense? [00:40:00] Yeah, Derrick, that's exactly right and in fact, cube root in this respect is stronger than the square root and the fourth root is stronger than that and those are all power transformations.
The log is basically stronger than all of them in that respect.
Yeah, I think usually that's most of it.
Sometimes some transformations are natural, for example, if you're dealing with volume of something or you're measuring the volume of something, often a cube root will do really well because maybe what's normally distributed [00:41:00] is the length and then everything else scales proportionately if you're dealing with area square root might work particularly well.
Sometimes you can actually think about what a particular effect means substantively and then get a good transformation from that.
If you have some multiplicative effect going on, then log often works really well.
Yeah, that's right.
I think different variables might call for different transformations and some for none at all.
So I wouldn't use a one-size-fits-all approach.
But you're on the right track.
Foster, it's a good question.
[00:42:00] There are several.
One thing unfortunately is usually the situation where you have more systematic ways to do it or when you have multiple groups that you're comparing and you want to, say, transform to equalise the variance within the groups.
Hopefully, I'm going to get this right.
[00:43:00] These are all power transformation like square roots and similar log.
Log you can view as an extreme power transformation but one systematic way to get those is using the Box-Cox transformation, which again, I'm not going to go through into much detail because I probably need to brush up as well.

But the idea there is you look at the relationship between group means and group variances and from that you derive the optimal transformation.
That's the suggestions I have.
I'll post my solutions which might not necessarily be the best possible solutions for that shortly.
Am I sharing the right screen? [00:44:00] I'm sorry, this is what I wanted to share.
This is the Box-Cox transformation which is a more systematic way to deal with this.
Again, it's mostly used in the context of regression where you have multiple groups and you want to transform such that within-group variances are the same or the mean-variance relationship is stable.
[00:45:00] Tarika, it's a good question, I don't remember the stuff off my head actually.
Apologies for that.
I think I don't know whether it passes for all of them.
I think it makes it past for some of them.
One thing and I think I also talked a bit about that in notes, the speed is out of the way, but it is an important point.
The thing is, you basically have this weird situation where the more power you have to [00:46:00] detect deviations from normality, the less you actually need the normality in the first place.
Or conversely, more precisely the bigger your sample size, the less you actually need underlying normality, but also, the more easily you can detect deviations from it.
The upshot of that is just that we shouldn't overlie in these tests.
Now, unfortunately, that does raise the question of, what good are they? What do we want? When can we assume normality? Probably the best way to do it is things like sensitivity analysis and methods like Bootstrap, where you want to basically see how much the normality is potentially hurting your conclusions.
[00:47:00] I don't know how much that helped.
Yes, I thought it was somewhere in the notes but anyway, yes, there is a multivariate central limit theorem.
It works exactly like the ordinary central limit theorem and no matter the original structure, the mean is going to be a multivariate normal.
We do approach multivariate normality.
[00:48:00] The trigger point B, they should put it in the next version of the notes.
I think the next question was also from you.
That was about Week 1, Topic 3.
It has very few exercises.
Yeah, I ended up dumping all the exercises into the next part because in and of themselves, point estimates are not very interesting, at least not if you're a statistician because we care about uncertainty.
[00:49:00] This Topic 4 is where we apply these ideas.
I think the last item was also within question or last item posted to the thread.
There is one question about Mardia's Skewness and Kurtosis.
Let's talk about that.
Switch modes here.
Fundamentally, when we [00:50:00] do statistical hypothesis tests or tests of statistical hypotheses.
What we ultimately do is we take some summary of the data and then figure out what we would expect from that summary of the data under the null hypothesis.
Then figure out whether what we observe, how far away is that from our expectations.
You can compute things like p-values or the certain measures of that.
But fundamentally that's what we do.
We ask, okay, under the null hypothesis, what would we expect to see and what is the range of things we'd expect to see.

Now, the thing we do see, how well does it turn there? In this case, the Mardia's measures are basically, they test for two [00:51:00] properties of normal distribution that we know very well.
The first one is skewness and the second one is kurtosis.
Now, in a univariate context, the skewness of the normal distribution is, what do you think? Actually yellow.
I think maybe I should define skewness.
Let me switch to the document camera and I'll define skewness for you.
The skewness of distribution for some random variable X, and this is just univariate for the moment.
[00:52:00] There's a number of ways of writing it but one is the expected value of X cubed minus the cube of the expected value of X.
Then typically we take that and we divide it by the variance of X to the power of three halves.
So this would be the normalised centre of the skewness.
It makes sense.
Sorry, I made a mistake.
I'm sorry.
I'm trying to do this from memory.
So [00:53:00] the expected value of X minus the expected value of X cubed, I'm sorry, yes.
This is the correct skewness.
For the quadratic moment is the same, for cube one it's not.
Sorry, this is the current one.
So we take the x, we centre it, and we take the cubes of these deviations from x.
Now, if we were to take the square then we would lose the direction but when we take a cube we don't.
Makes sense? So what would be the skewness of a normal distribution? It would be zero.
[00:54:00] In fact, symmetric distribution, you would expect this to be zero.
The skewness of these two cancel out.
Now, it turns out that, well, and so the idea is that if we take the sample skewness, so maybe something like that.
If you take the sample skewness, sorry.
One over n.
We would expect that if the distribution is really skewed then we would expect that value to be pretty far from zero.
We can also think about kurtosis.
Now, there's different ways to interpret it.
[00:55:00] It's the thickness of a tail of distribution because for example you can have a distribution that maybe looks like this and a distribution that looks like this.
You can tweak them until they have the same standard deviation.
But the idea is that one cuts off these observations very quickly whereas the other one stretches them out.
That's the difference between kurtosis.
This one would be said to have heavy tails and this one would be said to have no tails.
So kurtosis is a similar measure, except it's the fourth moment.
[00:56:00] It turns out that if you evaluate this on a normal distribution you will get a value of three.
In fact, sometimes when software packages report kurtosis they report this quantity minus three just to say that this is the kurtosis relative to the normal distribution.
The idea is that again, we can compute the kurtosis of distribution and if it's far away from three that's evidence that we don't have normality.
So this is the multivariate generalisation [00:57:00] of this.

So then what Mardia did is generalised that to multivariate distributions.
It's a little bit ad hoc but I guess they were able to show the property of this distribution, so kudos to them for that.
First you have the multivariate skewness.
Actually this one is hard to explain so I'm going to go talk about kurtosis first.
It's pretty straightforward.
One thing to keep in mind and this is when reading these is that expectation is a kind of a sum.
So the multiplication and powers take precedence over the sum so this expression should be read as the expectation of the square of [00:58:00] x minus Mu transpose sigma inverse x minus b.
So what we're doing here is we're basically taking this quantity which is also known as the Mahalanobis distance of X from its mean, but this is basically a measure of how far X is from its mean and its a squared measure of the distance.
Then we square it again.
So in some sense we now have the fourth power of the distance of X from its mean.
Then we take the expected value of that.
So that's analogous to taking that the numerator of the kurtosis.
Now, one last thing is because we have this Sigma inverse in here, that's where we divide by the variance, and because we square after we do that then we divide it by variance squared.
Make sense? In fact, you can show [00:59:00] that for a normal distribution with any mean and any variance, this is the true value of this quantity.
Now, for the skewness, it's a little bit funky.
The way it works is this.
If you just start with this quantity and take it to the three-halves power, well, you lose all information about skewness because you're squaring things in here.
But you don't want to square things.
You want to perhaps use this.
Perhaps what you want to do is you want to use some.
Have [01:00:00] your variable you're interested in on one side of the sigma inverse, and then another variable with the same distribution but independent of it on the other side of sigma inverse so that way, you still have that nice quadratic form.
But also what will no longer really quadratic.
You have that nice form.
But now it can be either positive or negative, and then you can cube it to get your skewness.
Now, obviously this is a theoretical quantity, and it turns out that basically if you take this expectation, well, you can see what happens.
If these are independent, you can show that at least for normal distribution the skewness is 0.
Now, the question is, what if you have data? Well, if you have data, then you don't really have separate X and Y, you just have one of these.
[01:01:00] What you do instead is something like this.
For kurtosis is pretty straight forward, essentially, you take the form, X minus X bar sample variance inverse X minus X bar, and you square that.
That's your fourth moment.
Whereas for skewness, you have to be a bit more clever because those Ys have to come from somewhere, and the idea is that if we have an independent sample, well, we have a value X_j same, and then also we have added value X_i, [01:02:00] and a value of X_j would be independent of the X_i, but it will come from the same distribution presumably.
Maybe what we only have to do is just scramble these values in order to get an approximation of this equation, 1.

6.

It turns out that at least for large enough samples, you can show that we can use these sample quantities.

We can plug your sample quantities that we can actually derive their approximate distribution, and they turn out to be chi-squared [01:03:00] and standard normal respectively.

The idea is that the way the Mardia test works is that what you will do is you would compute these quantities and I will do that for you most of the time, and then you would compare them.

You would compute the p-value based on the chi-squared distribution, and essentially, the bigger the K_1 here, the more likely you have some skewness and figure the K_2 the smaller, the more likely you'll get some kurtosis.

That's of how Mardia works.

Again, one thing we can do as well with kurtosis here is that [01:04:00] because we have this measure of these deviations from the mean, there are then potentially taken the fourth power.

That makes the outliers really visible.

That's the Mardia's skewness and kurtosis in the software demo, actually interesting with the first version of this course they used one set of diagnostics by default.

This is Mardia by default.

Whereas in a more recent version they switched to using a different test, and I'm sorry, I still haven't decided which one I like better.

[01:05:00] Well, I hope that helps.

Are there any other questions? Oh, yes.

[01:06:00] This notation is actually based on the vertical line use when talking about conditional probabilities.

You can think of this as the standard deviation or variance matrix of the first set of variables, conditional on the second set of variables, and in fact, I think if you go back a bit.

[01:07:00] If we're looking at this, you might notice that this expression a bit familiar, sigma 11 minus sigma 12, sigma 22 inverse sigma 21.

Whereas the other one is essentially the same form.

These are variances sample and population of variables at one conditional variable set 2.

That's a good question.

[01:08:00] In order for everything to fit together, in order for the matrices to be conformable, you would have to be the same dimension as S11.

So in this case, S11 is an r by r matrix, so it is smaller, and S22 would be n p minus r by p minus r matrix.

[01:09:00] Yeah, I think I apologise, I haven't been very clear on that.

Quiz 1 will not be timed.

Quiz 1, from its opening to its closing, there's no time limit.

That will take very long.

Amy, [01:10:00] yes.

The way Moodle quizzes work is that you can leave a quiz and it'll save your answers and will navigate away from the quiz and to save your answers, and then you can come back to it.

For the sake of my health, I only need to get up and pace for a little bit.

Well, let's take a five-minute break and I'll be back shortly.

[01:11:00] [01:12:00] [01:13:00] [01:14:00] [01:15:00] [01:16:00] [01:17:00] [01:18:00] All right, turn back.

Is there any more questions? That case I'm going to switch into consultation mode so that just means I'm going to mute myself and turn off the camera, [01:19:00] but I'm still watching the chat.

I'll actually just need to double-check that the instructions are correct and then I'll put it back probably by the end of this week.

Sorry about that.
I just needed to double-check that nothing has changed.
The dates and such are still correct.
Is there any other questions I can help? Alex, yes.
Basically, the website was misconfigured and so it sent everybody including me to the last year's instance of blackboard collaborate.
That is the one [01:20:00] for last year's instance of this class, which is why nobody could get in except for me because I was the instructor then as well so I could get in.
But then if you could because you were enrolled now and not a year ago.
In the end, I ended up posting guest links and so eventually we had a session.
The problem is that now nobody can access the recordings except for me.
What I did is I downloaded the recording and I put a link up to it.
I don't know whether you can use that, but it should be there.
Now, the website judging by the fact that we're all here, the current think is in fact correct.
Oh, excellent.
[inaudible 01:20:48] [01:21:00] I can link you to the post.
Let me try a direct link to it.
That is something that's been popping up randomly on add pages.
I think they have some glitch on the website.
We've been playing whack-a-mole with those things.
If you see that problem, please let me know and I will get it fixed but right now, basically, one of the technicians goes in and does some voodoo that seems to work, but it's not actually clear what's going on.
[01:22:00] He has done that suddenly it happened.
Oh, so it's not just us.
I'm not sure whether I should be happy it's not just us or sad that they are still haven't fixed it.
Any other questions? If not, I'm going to go into hibernation.
If you have a question, either raise your hand or type something into chat.
[01:23:00] [01:24:00] [01:25:00] [01:26:00] [01:27:00] [01:28:00] [01:29:00] [01:30:00] [01:31:00] [01:32:00] [01:33:00] [01:34:00] [01:35:00] [01:36:00] [01:37:00] [01:38:00] [01:39:00] [01:40:00] [01:41:00] [01:42:00] [01:43:00] [01:44:00] [01:45:00] [01:46:00] [01:47:00] All right.
Well, I guess that's it.