

As we have already discussed, we can estimate an unknown mean of a certain random variable by generating several independent samples of that random variable and taking their average. And this procedure is well justified, because of the weak law of large numbers, which tells us that this estimator converges when  $n$  goes to infinity, in probability, to the true mean. Now we can apply this idea more generally.

Suppose we want to estimate the expected value of a function of a random variable  $X$ . Now  $g$  of  $X$  is itself a random variable. So if we have samples of  $g$  of  $X$ , we can use the same procedure. How do we do that? We generate independent samples of  $X$ , and these give us independent samples of  $g$  of  $X$ . We use those independent samples, we average them, and by the weak law of large numbers, this quantity, as  $n$  goes to infinity, will converge in probability to the expected value of  $g$  of  $X$ .

We already used an idea of this form when we tried to estimate an unknown variance. A variance is defined as an expectation. And now we can generate samples of  $X$ , many independent samples, calculate this quantity, and take the average. However, we might not know the mean of the distribution. So instead of the true mean, we use an estimated mean, which is estimated the usual way using a sample average.

So when  $n$  is large, this estimated mean is close to the true mean. So using the estimated mean here will not make a substantial difference. And then we have essentially independent samples of this quantity. And by averaging them, we obtain an estimate of the variance, which asymptotically, as  $n$  goes to infinity, will be equal to the true variance.

Now we can push this idea even further. Suppose we wish to estimate a covariance. What's a natural way of doing this? We can generate independent samples of the pair of the random variables  $X$  and  $Y$ , so this will be a typical independent sample, and replace the expected value by a sample average. That is, we take our  $i$ -th sample,  $i$ -th pair, and calculate this quantity, which looks very much like the quantity in here except that we're using the estimated means in place of the true means.

We obtain these quantities and average  $n$  of them, again using the weak law of large numbers. One can argue that this estimate will converge to the true value of the covariance as  $n$  goes to infinity. And once we have estimates of a covariance and of variance, then we can use that to estimate correlation

coefficients.

Look at this formula, which is the definition of the correlation coefficient. If we just replace all quantities involved here by corresponding estimates, this gives us an estimate of the correlation coefficient. All of these ways of forming estimates can be shown. We are omitting the details of the argument, but hopefully you get the idea by now.

All of these quantities are consistent estimators. That is, when the sample size goes to infinity, they approach the correct values of what we're trying to estimate. So this is just an opening of what else a statistician might be interested in. And if you're wondering what's the further agenda after this point, it would be something like the following.

Typically, a statistician might want to say as much as possible about the probability distribution of an estimator. For example, we have here an estimate of a covariance. This estimate is going to be a random variable because it is determined by random quantities. What is the probability distribution of this quantity? Can we describe it approximately? What is the mean squared error associated with this estimator?

And if you wish to construct confidence intervals how would you do it? These are all topics that statisticians have studied in depth, and you could see more about these topics if you were to take a further class on statistics and inference. But we will not go any deeper in this course.