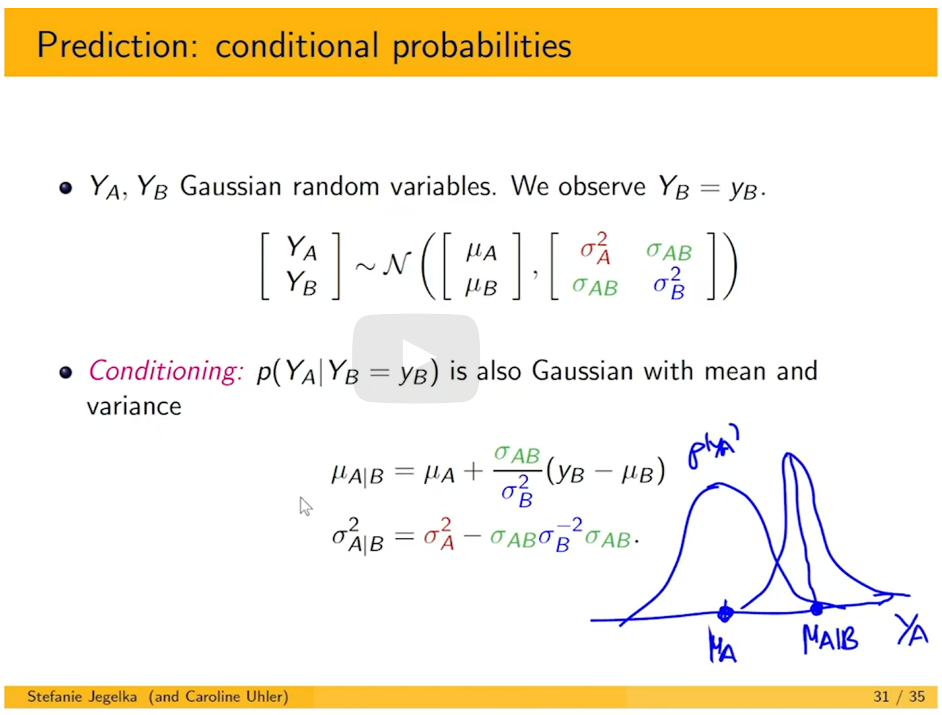
5. Conditional Probability

Bookmark this page

Exercises due May 19, 2021 19:59 EDT

Conditional Probability



10:00 / 10:16 1.50x

So I actually get them always to be fairly close to each other

if they are highly correlated.

But all of this, essentially, is just

expressed in these two equations.

And they will be of central importance

going forward when we start with more modeling.

And, moreover, they will be very important

also for the generalization to the matrix case

that we'll look at next.

Video

Download video file

Transcripts

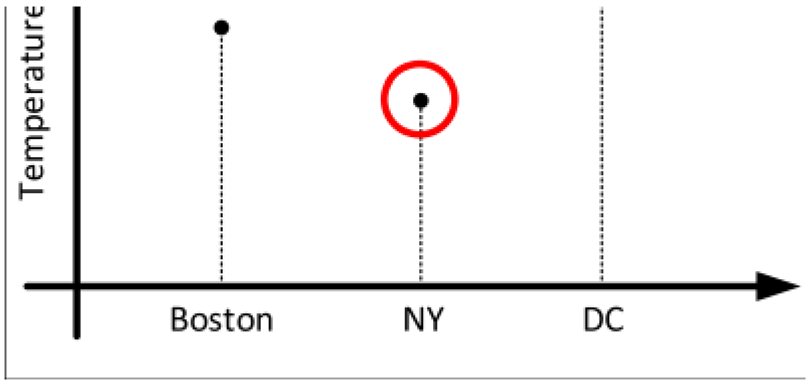
Download SubRip (.srt) file

Download Text (.txt) file

Definition 5.1 Two random variables are defined as jointly Gaussian if their joint distribution is a multivariate Gaussian random variable. For example, for two Gaussian random variables X_1 and X_2, one can define a new two-dimensional random variable $Y = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$. The resulting probability density function for the joint distribution can be defined as

$$p\left(\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \mid \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}\right) = \frac{1}{2\pi(\sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21})^{1/2}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)^T \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}^{-1} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)\right)$$

Conditional Distribution of Gaussian Random Variables



31: Temperature values at three different cities.

Recall the example shown in the above figure, and for simplicity consider only two random variables. Thus, assume there are two cities, *City 1* and *City 2*. Moreover, assume their respective temperatures behave like Normal random variables, denoted as X_1 and X_2, and we have sufficient information to construct good estimates of their respective statistics, i.e., (μ_1, μ_2) and their covariance matrix

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

are known.

Now, assume we have access to a measurement of the temperature in the *City 2*, and we are interested in estimating the temperature in *City 1*. Let us denote the temperature in *City 2* as x_2.

Recall that we know that the temperature in *City 1* follows a Normal distribution with mean μ_1 and a standard deviation of σ_1. So, in the absence of any other information, one could expect that μ_1 is a reasonable estimate for the temperature in *City 1*. We also have additional information: we know X_1 is correlated with X_2, and we have observed that $X_2 = x_2$. Can we get a better estimate for the temperature in *City 1*? A better estimate, in this case, refers to the reduction of the variance on the estimated value, which initially is denoted as σ_1.

In order to do so, we will need to compute the conditional probability of X_1 given $X_2 = x_2$, under the assumption that X_1 and X_2 are correlated and jointly Gaussian.

Initially, recall that the conditional density function for X_1 given X_2 is defined as

$$\begin{aligned} p_{X_1|X_2=x_2}(x) &= \frac{p_{X_1,X_2}(x, x_2)}{p_{X_2}(x_2)} \\ &\propto \frac{\exp\left(-\frac{1}{2}\left(\begin{bmatrix} x-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right)^{\intercal}\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}^{-1}\left(\begin{bmatrix} x-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\right)\right)}{\exp\left(-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)} \\ & \text{Left as exersice...} \\ &\propto \exp\left(-\frac{1}{2}\left(\frac{\left[x-\left(\mu_1+\frac{\sigma_{12}}{\sigma_2^2}(x_2-\mu_2)\right)\right]^2}{\sigma_1^2-\frac{\sigma_{12}^2}{\sigma_2^2}}\right)\right). \end{aligned}$$

The result above implies that the conditional distribution $P(X_1 \mid X_2 = x_2)$, is still a Normal distribution with a new mean and a new standard deviation defined as

$$\mu_{X_1|X_2=x_2} = \mu_1 + \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_2^2}(x_2 - \mu_2) \tag{7.2}$$

$$\sigma^2_{X_1|X_2=x_2} = \sigma_1^2 - \frac{\mathrm{Cov}(X_1, X_2)^2}{\sigma_2^2}. \tag{7.3}$$

The result in these equations suggests a number of interesting interpretations.

- The conditional mean of X_1 given X_2 is the original mean estimate μ_1 shifted by a weighted version of the difference between the observation x_2 and the mean μ_2. Moreover, the weight is proportional to the ratio between the covariance $\mathrm{Cov}(X_1, X_2)$ between X_1 and X_2 and the variance of X_2.

- If X_1 and X_2 are independent, then Cov $(X_1, X_2) = 0$, therefore, observations of x_2 will not affect the original statistics $\mu_{X_1|X_2=x_2} = \mu_1$ and $\sigma^2_{X_1|X_2=x_2} = \sigma^2_1$.
- The conditional variance of X_1 given X_2 is the original variance, σ^2_1, minus the ratio between the covariance squared, Cov$(X_1, X_2)^2$, and the variance, σ^2_2.

Properties of Multivariate Gaussian Distribution

1 point possible (graded)

Is it possible that $\sigma_1 < \frac{\text{Cov}(X_1,X_2)}{\sigma_2}$?

- Yes
- No

Submit You have used 0 of 1 attempt

Conditional Probability

1 point possible (graded)

Assume the temperature in *City 1* X_1 is a Gaussian random variable with mean $\mu_1 = 60$ and $\sigma_1 = 10$, and that the temperature of *City 2* X_2 is is a Gaussian random variable with mean $\mu_2 = 90$ and $\sigma_2 = 20$. Moreover, we know that the covariance between X_1 and X_2 is 100. Today, we have observed that the temperature in *City 2* is 75. What is the probability that the new temperature in the *City 1* is bigger than 56.25?

Submit You have used 0 of 4 attempts

Discussion

Hide Discussion

Topic: Module 5: Environmental Data and Gaussian Processes:Environmental Data and Gaussian Processes / 5. Conditional Probability

Add a Post

Show all posts by recent activity

[STAFF]: typo in Defn 5.1? 1

[Staff] Potential error in the solution 4
There is a ??? in the solution of Properties of Multivariate Gaussian Distribution. Should there be some words there? Thank you very ...

< Previous Next >