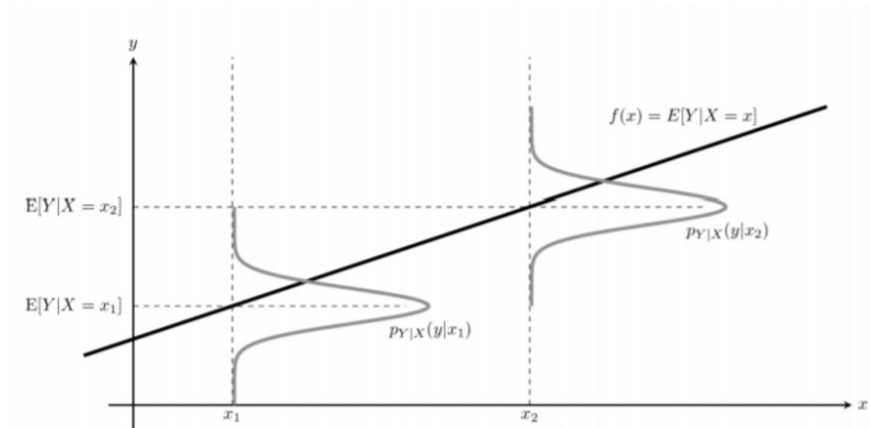
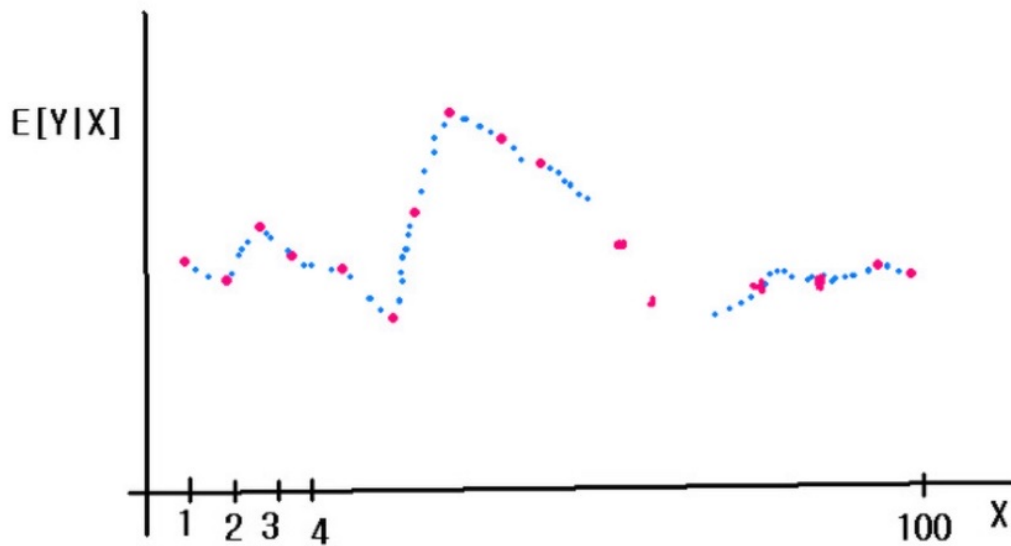


그렇다면 기댓값과 조건부 기댓값의 차이점은 무엇인가?



위와 같이 x_1, x_2 일때의 조건부 확률분포가 존재할때 각 확률분포의 평균을 찾아주는 것이 조건부기댓값이라고 부를 수 있다.



이 그림을 통해 $E[Y|X]$ 는 X 의 어떤 함수라는 것을 알 수 있습니다. X 가 확률변수이기 때문에 X 의 함수 $h(X)$ 도 확률변수이고 따라서 $E[Y|X]$ 도 확률변수입니다.

그럼 앞의 설명을 통해서 우리는 다음과 같은 것을 추측할 수 있습니다. 어떤 현상 Y 가 있고, Y 에 대해 설명해줄 수 있는, 또는 정보 X 가 있습니다. 우리는 X 의 정보를 통해 Y 를 최대한 잘 표현하거나, 설명하거나, 예측하고 싶다는 것이죠. 수학적으로 이야기해서 X 의 어떤 함수, $h(X)$ 를 통해 Y 를 잘 표현, 설명, 예측하고 싶다는 것이죠. 그럼 X 의 어떤 함수가 가장 잘 Y 를 표현, 설명, 예측해 줄 수 있을까요?

이미 답을 예상했겠죠. Y 를 가장 잘 설명, 표현, 예측하는 X 의 함수, $h(X)$ 는 $E[Y|X]$ 입니다. 가장 대표적인 것이 회귀분석에서 $h(X)=E[Y|X]=a+bX$ 이죠.