Conditional versus Marginal Distributions

Joint, Marginal, Conditional

- **A (scalar) random variable X has a distribution that we call the "distribution of X".**
- **Suppose we have two random variables; X and Y. Then there is the "joint distribution of X and Y", or the "joint distribution of (X, Y).**
- **The "distribution of X" still has the same meaning. But now there is also the "conditional distribution of X given Y". To emphasize the distinction we refer to the "distribution of X" as the "marginal distribution of X".**
- **Similarly, there is the "marginal distribution of Y" and the "conditional distribution of Y given X".**
- **The marginal distribution of Y and the conditional distribution of Y given X are identical if and only if X and Y are independent.**

Example 1

- $X \sim$ **Bernoulli(α). E**$[X] = P(X = 1) = 1 - P(X = 0) = \alpha$**.**
- Y **given** X **is distributed as Bernoulli(**$\beta_1 + \beta_2 X$**).**
- **Then the marginal distribution of** Y **is Bernoulli(**$\beta_1 + \beta_2 \alpha$**).** $P(Y = 1) = \beta_1 + \beta_2 \alpha$**.**
- **The joint distribution of** (X, Y) **is specified as a list or a prescription of how to compute** $P(X = x, Y = y)$ **for all four possible** (x, y) **pairs.**
- **E**$[XY] = P(XY = 1) = P(X = 1, Y = 1) = P(X = Y = 1) = P(X = 1)P(Y = 1|X = 1) = \alpha(\beta_1 + \beta_2)$**.**
- **E**$[X|Y = 1] = P(X = 1|Y = 1) = P(X = 1, Y = 1)/P(Y = 1) = \alpha(\beta_1 + \beta_2)/(\beta_1 + \beta_2 \alpha)$**, not equal to** α **unless** $\beta_2 = 0$ **or** $\alpha = 0$**.**
- **Exercise: E**$[X|Y = 0] = ?$**.**

Example 2

- **Variance reduction by conditioning?**

  Take $P(X = 1) = 0.6, P(Y = 1) = 0.2, P(XY = 1) = 0.1$**.**

- **var**$(Y|X = 0) = 0.1875 >$ **var**$(Y) = 0.16$ **(inflation)**

- **var**$(Y|X = 1) \approx 0.139 <$ **var**$(Y) = 0.16$ **(reduction)**

- **Compare: In the bivariate normal distribution, var**$(Y|X) =$ **var**$(Y)(1-\rho^2)$**.**

  Variance reduction by conditioning unless the correlation $\rho = 0$**.**

  Another feature of the bivariate normal: var$(Y|X)$ **does not depend on** X**.**

Example 3

- $(X,Y) \sim$ **bivariate normal with** $\mathbf{E}[X] = \mu_1, \mathbf{E}[Y] = \mu_2, \mathbf{var}(X) = \sigma_{11}, \mathbf{var}(Y) = \sigma_{22}, \mathbf{cov}(X,Y) = \sigma_{12} = \sigma_{21}$**.**
  $\rho := \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$
- $X \sim N(\mu_1, \sigma_{11})$**,** $Y \sim N(\mu_2, \sigma_{22})$**.**
- X **given** Y **is** $\sim$

$$N\left(\mu_1 + \sigma_{12}\sigma_{22}^{-1}(Y - \mu_2), \sigma_{11} - \sigma_{12}\sigma_{22}^{-1}\sigma_{21}\right)$$

- $\mathbf{var}(X|Y) = \sigma_{11}(1-\rho^2))$**.**
- X **and** Y **vectors,** $(X,Y) \sim$ **multivariate normal?**