ECE 361 Probability for Engineers (Fall, 2016)
Lecture 9b

§4.2 Covariance and correlation

Example. (CONTINUED FROM PREVIOUS LECTURE). Let the bins selected for the n balls be $B_1, \ldots, B_n$, where each $B_t \in [k]$ for $t \in [n]$. Observe we can write $X_i = \sum_{t=1}^n \mathbf{1}(B_t = i)$ and likewise $X_j = \sum_{t=1}^n \mathbf{1}(B_t = j)$. Here, $\mathbf{1}(A)$ is an *indicator* RV, with

$$\mathbf{1}(A) = \begin{cases} 1, & A \text{ true} \\ 0, & A \text{ false} \end{cases} \tag{1}$$

Observe that $\mathbf{1}(A)$ is a Bernoulli RV with probability of success $\mathbb{P}(A)$, and thus it has mean $\mathbb{E}[\mathbf{1}(A)] = \mathbb{P}(A)$ and variance $\mathbb{P}(A)(1-\mathbb{P}(A))$. In our setting, $\mathbf{1}(B_t = i)$ equals one if $B_t = i$ and zero else, and has expectation $\mathbb{E}[\mathbf{1}(B_t = i)] = \mathbb{P}(B_t = i) = p_i$.

It follows that

$$\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}\left(\sum_{t=1}^n \mathbf{1}(B_t = i), \sum_{t=1}^n \mathbf{1}(B_t = j)\right) = \sum_{s=1}^n \sum_{t=1}^n \mathrm{Cov}(\mathbf{1}(B_s = i), \mathbf{1}(B_t = j)). \tag{2}$$

There are two types of terms in this sum: *i*) $s = t$ and *ii*) $s \neq t$. First consider the case $s = t$, in which case we compute

$$\begin{aligned} \mathrm{Cov}(\mathbf{1}(B_s = i), \mathbf{1}(B_s = j)) &= \mathbb{E}[(\mathbf{1}(B_s = i) - p_i)(\mathbf{1}(B_s = j) - p_j)] \\ &= \sum_{i'=1}^k (\mathbf{1}(i' = i) - p_i)(\mathbf{1}(i' = j) - p_j) p_{i'} \\ &= (1 - p_i)(0 - p_j)p_i + (0 - p_i)(1 - p_j)p_j + (0 - p_i)(0 - p_j)(1 - p_i - p_j) \\ &= -(1 - p_i)p_i p_j - (1 - p_j)p_i p_j + p_i p_j (1 - p_i - p_j) \\ &= -p_i p_j + p_i^2 p_j - p_i p_j + p_i p_j^2 + p_i p_j - p_i^2 p_j - p_i p_j^2 \\ &= -p_i p_j \end{aligned} \tag{3}$$

Next consider the case $s \neq t$, in which case we compute

$$\begin{aligned} \mathrm{Cov}(\mathbf{1}(B_s = i), \mathbf{1}(B_t = j)) &= \mathbb{E}[(\mathbf{1}(B_s = i) - p_i)(\mathbf{1}(B_t = j) - p_j)] \\ &= \mathbb{E}[\mathbf{1}(B_s = i) - p_i]\mathbb{E}[\mathbf{1}(B_t = j) - p_j] = 0 \times 0 = 0 \end{aligned} \tag{4}$$

We now substitute these expressions into the earlier expression:

$$\begin{aligned} \mathrm{Cov}(X_i, X_j) &= \sum_{s=1}^n \sum_{t=1}^n \mathrm{Cov}(\mathbf{1}(B_s = i), \mathbf{1}(B_t = j)) \\ &= \sum_{s=1}^n (-p_i p_j) = -n p_i p_j. \end{aligned} \tag{5}$$

Finally, we compute the correlation:

$$\begin{aligned} \rho(X_i, X_j) &= \frac{\mathrm{Cov}(X_i, X_j)}{\mathrm{Std}(X_i)\mathrm{Std}(X_j)} \\ &= -\frac{n p_i p_j}{\sqrt{n p_i (1 - p_i) n p_j (1 - p_j)}} \\ &= -\sqrt{\frac{p_i}{1 - p_i} \times \frac{p_j}{1 - p_j}} \end{aligned} \tag{6}$$

Example. Let (U, V, W) be independent continuous RVs, each uniformly distributed over $[0, 1]$. Define the pair of continuous RVs (X, Y) with $X = U/(U + W)$ and $Y = V/(V + W)$. Find the correlation of (X, Y).

We first find the expected values of (X,Y):

$$\mathbb{E}[X] = \mathbb{E}\left[\frac{U}{U+W}\right] = \int_0^1 \int_0^1 \frac{u}{u+w} \mathrm{d}w\mathrm{d}u = \int_0^1 u \log(1+1/u)\mathrm{d}u = \frac{1}{2}\left(u + u^2 \log(1+1/u) - \log(1+u)\right)\Big|_0^1 = \frac{1}{2} \quad (7)$$

and thus $\mathbb{E}[Y] = 1/2$ as well. We next find the covariance:

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}[(X-1/2)(Y-1/2)] \\
&= \mathbb{E}[XY - X/2 - Y/2 + 1/4] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]/2 - \mathbb{E}[Y]/2 + 1/4 \\
&= \mathbb{E}[XY] - 1/4 \\
&= \int_0^1 \int_0^1 \int_0^1 \frac{u}{u+w} \times \frac{v}{v+w} \mathrm{d}u\mathrm{d}v\mathrm{d}w - \frac{1}{4} \\
&\vdots \\
&= \frac{1}{18}(6 + \pi^2 - 12(2-\log 2)\log 2) - \frac{1}{4} \approx 0.02775.
\end{aligned} \quad (8)$$

We next find the variances by first finding the expected squares:

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}\left[\left(\frac{U}{U+W}\right)^2\right] = \int_0^1 \int_0^1 \left(\frac{u}{u+w}\right)^2 \mathrm{d}u\mathrm{d}w = 1 - \log 2 \\
\mathbb{E}[Y^2] &= \mathbb{E}\left[\left(\frac{V}{V+W}\right)^2\right] = \int_0^1 \int_0^1 \left(\frac{v}{v+w}\right)^2 \mathrm{d}v\mathrm{d}w = 1 - \log 2
\end{aligned} \quad (9)$$

and then substituting:

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1 - \log 2 - \frac{1}{4} = \frac{3}{4} - \log 2 \\
\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 1 - \log 2 - \frac{1}{4} = \frac{3}{4} - \log 2
\end{aligned} \quad (10)$$

Finally, we obtain the correlation:

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Std}(X)\mathrm{Std}(Y)} = \frac{\frac{1}{18}(6+\pi^2 - 12(2-\log 2)\log 2) - \frac{1}{4}}{\frac{3}{4} - \log 2} \approx 0.48811. \quad (11)$$

Example. Let (U,V,W) be independent and uniformly distributed RVs and define $X = UV$ and $Y = (1-U)W$. Find the correlation of (X,Y). We first find the covariance:

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= \mathbb{E}[UV(1-U)W] - \mathbb{E}[UV]\mathbb{E}[(1-U)W] \\
&= \mathbb{E}[U(1-U)]\mathbb{E}[V]\mathbb{E}[W] - \mathbb{E}[U]\mathbb{E}[V]\mathbb{E}[(1-U)]\mathbb{E}[W] \\
&= \frac{1}{4}\mathbb{E}[U(1-U)] - \frac{1}{16}.
\end{aligned} \quad (12)$$

Next:

$$\begin{aligned}
\mathbb{E}[U(1-U)] &= \int_0^1 u(1-u)\mathrm{d}u \\
&- \int_0^1 u\mathrm{d}u - \int_0^1 u^2\mathrm{d}u \\
&= \frac{1}{2}u^2\Big|_0^1 - \frac{1}{3}u^3\Big|_0^1 = \frac{1}{6}.
\end{aligned} \quad (13)$$

Thus: $\mathrm{Cov}(X,Y) = -1/48$. Next:

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[UV] = \mathbb{E}[U]\mathbb{E}[V] = \frac{1}{4} \\
\mathbb{E}[Y] &= \mathbb{E}[(1-U)W] = \mathbb{E}[(1-U)]\mathbb{E}[W] = \frac{1}{4} \\
\mathbb{E}[X^2] &= \mathbb{E}[(UV)^2] = \mathbb{E}[U^2]\mathbb{E}[V^2] \\
\mathbb{E}[Y^2] &= \mathbb{E}[((1-U)W)^2] = \mathbb{E}[(1-U)^2]\mathbb{E}[W^2]
\end{aligned} \tag{14}$$

Compute:

$$\begin{aligned}
\mathbb{E}[U^2] = \mathbb{E}[V^2] = \mathbb{E}[W^2] &= \int_0^1 u^2 \mathrm{d}u = \frac{1}{3} \\
\mathbb{E}[(1-U)^2] &= \int_0^1 (1-u)^2 \mathrm{d}u = \frac{1}{3}
\end{aligned} \tag{15}$$

Substituting:

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{9} - \frac{1}{16} = \frac{7}{144} \tag{16}$$

and of course $\mathrm{Var}(Y) = \frac{7}{144}$ as well. Finally:

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Std}(X)\mathrm{Std}(Y)} = \frac{-1/48}{\sqrt{(7/144)(7/144)}} = -9/21 \approx -0.42857. \tag{17}$$

§5.1 Markov and Chebychev inequalities

The Markov inequality asserts: if an RV X takes only nonnegative values then

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}, \ \forall a > 0 \tag{18}$$

The proof is insightful. Define

$$Y_a = \left\{ \begin{array}{ll} 0, & X < a \\ a, & X \geq a \end{array} \right. , \tag{19}$$

and observe $Y_a \leq X$ for all $a > 0$. Then $\mathbb{E}[Y_a] \leq \mathbb{E}[X]$ and $\mathbb{E}[Y_a] = a\mathbb{P}(X \geq a)$.

The Chebychev inequality asserts: if X is a RV with mean μ and variance σ^2 then

$$\mathbb{P}(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}, \ c > 0. \tag{20}$$

Again the proof is insightful. Define the nonnegative RV $(X-\mu)^2$ and use $a = c^2$ and apply the Markov inequality:

$$\mathbb{P}(|X-\mu| \geq c) = \mathbb{P}((X-\mu)^2 \geq c^2) \leq \frac{\sigma^2}{c^2}. \tag{21}$$

The Chebychev inequality is often used to obtain an upper bound on the complementary CDF as follows:

$$\mathbb{P}(X > x) = \mathbb{P}(X - \mu > x - \mu) \leq \mathbb{P}(|X-\mu| > x - \mu) \leq \frac{\sigma^2}{(x-\mu)^2}. \tag{22}$$

References

[1] *Introduction to Probability, 2nd Edition* by Dimitri P. Bertsekas and John N. Tsitsiklis, Athina Scientific Press, 2008.