CMPT 210: Probability and Computing

Lecture 20

Sharan Vaswani

March 26, 2024

Recap

Variance: Standard way to measure the deviation from the mean. For r.v. X, $\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{x \in \text{Range}(X)} (x - \mu)^2 \Pr[X = x]$, where $\mu := \mathbb{E}[X]$.

Alternate Definition: $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Standard Deviation: For r.v. X, the standard deviation of X is defined as $\sigma_X := \sqrt{\text{Var}[X]} = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}$.

For constants a, b and r.v. R, $\text{Var}[aR + b] = a^2 \text{Var}[R]$.

Pairwise Independence: Random variables $R_1, R_2, R_3, \ldots R_n$ are pairwise independent if for any pair R_i and R_j, for $x \in \text{Range}(R_i)$ and $y \in \text{Range}(R_j)$, $\Pr[(R_i = x) \cap (R_j = y)] = \Pr[R_i = x] \Pr[R_j = y]$.

Linearity of variance for pairwise independent r.v's: If $R_1, \ldots, R_n$ are pairwise independent, $\text{Var}[R_1 + R_2 + \ldots R_n] = \sum_{i=1}^{n} \text{Var}[R_i]$.

Covariance

For two random variables R and S, the covariance between R and S is defined as:

$$\text{Cov}[R, S] := \mathbb{E}[(R - \mathbb{E}[R])\,(S - \mathbb{E}[S])] = \mathbb{E}[RS] - \mathbb{E}[R]\,\mathbb{E}[S]$$

$$\begin{aligned}
\text{Cov}[R, S] &= \mathbb{E}[(R - \mathbb{E}[R])\,(S - \mathbb{E}[S])] \\
&= \mathbb{E}\left[RS - R\,\mathbb{E}[S] - S\,\mathbb{E}[R] + \mathbb{E}[R]\,\mathbb{E}[S]\right] \\
&= \mathbb{E}[RS] - \mathbb{E}[R\,\mathbb{E}[S]] - \mathbb{E}[S\,\mathbb{E}[R]] + \mathbb{E}[R]\,\mathbb{E}[S] \\
\implies \text{Cov}[R, S] &= \mathbb{E}[RS] - \mathbb{E}[R]\,\mathbb{E}[S] - \mathbb{E}[S]\,\mathbb{E}[R] + \mathbb{E}[R]\,\mathbb{E}[S] = \mathbb{E}[RS] - \mathbb{E}[R]\,\mathbb{E}[S]
\end{aligned}$$

Covariance generalizes the notion of variance to multiple random variables.

$$\text{Cov}[R, R] = \mathbb{E}[R\,R] - \mathbb{E}[R]\,\mathbb{E}[R] = \text{Var}[R]$$

If R and S are independent r.v's, $\mathbb{E}[RS] = \mathbb{E}[R]\,\mathbb{E}[S]$ and $\text{Cov}[R, S] = 0$.

The covariance between two r.v's is symmetric i.e. $\text{Cov}[R, S] = \text{Cov}[S, R]$.

Covariance

For two arbitrary (not necessarily independent) r.v's, R and S,

$$\mathrm{Var}[R+S] = \mathrm{Var}[R] + \mathrm{Var}[S] + 2\,\mathrm{Cov}[R,S]$$

Recall from Lecture 19, Slide 6, where we showed that,

$$\mathrm{Var}[R+S] = \mathrm{Var}[R] + \mathrm{Var}[S] + 2(\mathbb{E}[RS] - \mathbb{E}[R]\,\mathbb{E}[S]) = \mathrm{Var}[R] + \mathrm{Var}[S] + 2\,\mathrm{Cov}[R,S].$$

If R and S are independent, $\mathrm{Cov}[R,S] = 0$ and we recover the formula for the sum of independent variables.

For $R = S$, $\mathrm{Var}[R+R] = \mathrm{Var}[R] + \mathrm{Var}[R] + 2\mathrm{Cov}[R,R] = \mathrm{Var}[R] + \mathrm{Var}[R] + 2\mathrm{Var}[R] = 4\mathrm{Var}[R]$ which is consistent with our previous formula that $\mathrm{Var}[2R] = 2^2\mathrm{Var}[R]$.

Generalization to multiple random variables $R_1, R_2, \ldots R_n$ (Recall from Lecture 19, Slide 7):

$$\mathrm{Var}\left[\sum_{i=1}^{n} R_i\right] = \sum_{i=1}^{n} \mathrm{Var}[R_i] + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}[R_i, R_j]$$

Covariance - Example

Q: If X and Y are indicator r.v's for events A and B respectively, calculate the covariance between X and Y

We know that $\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$. Note that $X = \mathcal{I}_A$ and $Y = \mathcal{I}_B$. We can conclude that $XY = \mathcal{I}_{A \cap B}$ since $XY = 1$ iff both events A and B happen.

$$\implies \mathbb{E}[X] = \Pr[A]\ ; \mathbb{E}[Y] = \Pr[B]; \mathbb{E}[XY] = \Pr[A \cap B]$$

$$\implies \mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = \Pr[A \cap B] - \Pr[A]\Pr[B]$$

If $\mathrm{Cov}[X, Y] > 0 \implies \Pr[A \cap B] > \Pr[A]\Pr[B]$. Hence,

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} > \frac{\Pr[A]\Pr[B]}{\Pr[B]} = \Pr[A]$$

If $Cov[X, Y] > 0$, it implies that $\Pr[A|B] > \Pr[A]$ and hence, the probability that event A happens increases if B is going to happen/has happened. Similarly, if $Cov[X, Y] < 0$, $\Pr[A|B] < \Pr[A]$. In this case, if B happens, then the probability of event A decreases.

Correlation

The correlation between two r.v's R_1 and R_2 is defined as:

$$\mathrm{Corr}[R_1, R_2] = \frac{\mathrm{Cov}[R_1, R_2]}{\sqrt{\mathrm{Var}[R_1]\,\mathrm{Var}[R_2]}}$$

$\mathrm{Corr}[R_1, R_2] \in [-1, 1]$ and indicates the strength of the relationship between R_1 and R_2.

If $\mathrm{Corr}[R_1, R_2] > 0$, then R_1 and R_2 are said to be positively correlated, else if $\mathrm{Corr}[R_1, R_2] < 0$, the r.v's are negatively correlated.

If $R_1 = R_2 = R$, then, $\mathrm{Corr}[R, R] = \frac{\mathrm{Cov}[R,R]}{\sqrt{\mathrm{Var}[R]\,\mathrm{Var}[R]}} = \frac{\mathrm{Var}[R]}{\mathrm{Var}[R]} = 1$.

If R_1 and R_2 are independent, $\mathrm{Cov}[R_1, R_2] = 0$ and $\mathrm{Corr}[R_1, R_2] = 0$.

If $R_1 = -R_2 = R$, then,

$$\mathrm{Corr}[R, -R] = \frac{\mathrm{Cov}[R, -R]}{\sqrt{\mathrm{Var}[R]\,\mathrm{Var}[-R]}} = \frac{\mathrm{Cov}[R, -R]}{\sqrt{\mathrm{Var}[R]\,(-1)^2\mathrm{Var}[R]}} = \frac{\mathrm{Cov}[R, -R]}{\mathrm{Var}[R]}$$
$$= \frac{\mathbb{E}[-R^2] - \mathbb{E}[R]\,\mathbb{E}[-R]}{\mathrm{Var}[R]} = \frac{-\mathbb{E}[R^2] + \mathbb{E}[R]\,\mathbb{E}[R]}{\mathrm{Var}[R]} = \frac{-\mathrm{Var}[R]}{\mathrm{Var}[R]} = -1$$

Questions?

Tail inequalities

Variance gives us one way to measure how "spread" the distribution is.

Tail inequalities bound the probability that the r.v. takes a value much different from its mean.

Example: Consider a r.v. X that can take on only non-negative values and $\mathbb{E}[X] = 99.99$. Show that $\Pr[X \geq 300] \leq \frac{1}{3}$.

$$\begin{aligned} \textit{Proof}: \mathbb{E}[X] &= \sum_{x \in \text{Range}(X)} x \Pr[X = x] = \sum_{x | x \geq 300} x \Pr[X = x] + \sum_{x | 0 \leq x < 300} x \Pr[X = x] \\ &\geq \sum_{x | x \geq 300} (300) \Pr[X = x] + \sum_{x | 0 \leq x < 300} x \Pr[X = x] \\ &= (300) \Pr[X \geq 300] + \sum_{x | 0 \leq x < 300} x \Pr[X = x] \end{aligned}$$

If $\Pr[X \geq 300] > \frac{1}{3}$, then, $\mathbb{E}[X] > (300)\,\frac{1}{3} + \sum_{x|0 \leq x < 300} x \Pr[X = x] > 100$ (since the second term is always non-negative). Hence, if $\Pr[X \geq 300] > \frac{1}{3}$, $\mathbb{E}[X] > 100$ which is a contradiction since $\mathbb{E}[X] = 99.99$.