CMPT 210: Probability and Computing

Lecture 20

Sharan Vaswani

November 19, 2024

Recap

- **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 $y \in \text{Range}(R_j)$, $\Pr[(R_i = x) \cap (R_j = y)] = \Pr[R_i = x] \Pr[R_j = y]$.
- **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**: $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.
- If $X \sim \text{Ber}(p)$, $\text{Var}[X] = p(1-p) \leq \frac{1}{4}$.

Back to throwing dice

Q: For a standard dice, if X is the r.v. equal to the number that comes up, compute $\text{Var}[X]$.

Recall that, for a standard dice, $X \sim \text{Uniform}(\{1,2,3,4,5,6\})$ and hence,

$$\mathbb{E}[X^2] = \sum_{x \in \{1,2,3,4,5,6\}} x^2 \Pr[X = x] = \frac{1}{6}\left[1^2 + 2^2 + \ldots + 6^2\right] = \frac{91}{6}$$

$$(\mathbb{E}[X])^2 = \left(\sum_{x \in \{1,2,3,4,5,6\}} x \Pr[X = x]\right)^2 = \left(\frac{1}{6}\left[1 + 2 + \ldots + 6\right]\right)^2 = \frac{49}{4}$$

$$\implies \text{Var}[X] = \frac{91}{6} - \frac{49}{4} \approx 2.917$$

Q: If $X \sim \text{Uniform}(\{v_1, v_2, \ldots v_n\})$, compute $\text{Var}[X]$.

$$\mathbb{E}[X] = \sum_{i=1}^{n} v_i \Pr[X = v_i] = \frac{1}{n}\left[v_1 + v_2 + \ldots v_n\right] \quad ; \quad \mathbb{E}[X^2] = \frac{1}{n}\left[v_1^2 + v_2^2 + \ldots v_n^2\right].$$

$$\implies \text{Var}[X] = \frac{\left[v_1^2 + v_2^2 + \ldots v_n^2\right]}{n} - \left(\frac{\left[v_1 + v_2 + \ldots v_n\right]}{n}\right)^2$$

Variance - Examples

Q: Calculate Var[W], Var[Y] and Var[Z] whose PDF's are given as:

$$\begin{aligned} W &= 0 && \text{(with } p = 1) \\ Y &= -1 && \text{(with } p = 1/2) \\ &= +1 && \text{(with } p = 1/2) \\ Z &= -1000 && \text{(with } p = 1/2) \\ &= +1000 && \text{(with } p = 1/2) \end{aligned}$$

Recall that $\mathbb{E}[W] = \mathbb{E}[Y] = \mathbb{E}[Z] = 0$.

$\text{Var}[W] = \mathbb{E}[W^2] - (\mathbb{E}[W])^2 = \mathbb{E}[W^2] = \sum_{w \in \text{Range}(W)} w^2 \Pr[W = w] = 0^2(1) = 0$. The variance of W is zero because it can only take one value and the r.v. does not "vary".

$\text{Var}[Y] = \mathbb{E}[Y^2] = \sum_{y \in \text{Range}(Y)} y^2 \Pr[Y = y] = (-1)^2(1/2) + (1)^2(1/2) = 1$.

$\text{Var}[Z] = \mathbb{E}[Z^2] = \sum_{z \in \text{Range}(Z)} z^2 \Pr[Z = z] = (-1000)^2(1/2) + (1000)^2(1/2) = 10^6$.

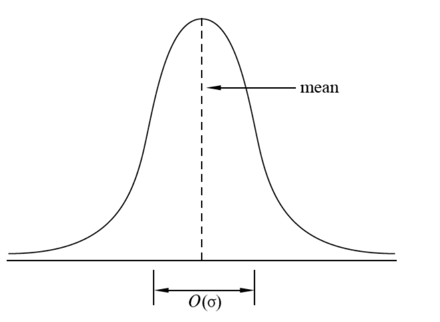
- Hence, the variance can be used to distinguish between r.v.'s that have the same mean.

Standard Deviation

Standard Deviation: For r.v. X, the standard deviation in X is defined as:

$$\sigma_X := \sqrt{\mathrm{Var}[X]} = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}$$

- Standard deviation has the same units as expectation.



- Standard deviation for a "bell"-shaped distribution indicates how wide the "main part" of the distribution is.

Variance - Examples

Q: If $R \sim \text{Geo}(p)$, calculate $\text{Var}[R]$.

$$\text{Var}[R] = \mathbb{E}[R^2] - (\mathbb{E}[R])^2 = \mathbb{E}[R^2] - \frac{1}{p^2}$$

Recall that for a coin s.t. $\Pr[\text{heads}] = p$, R is the r.v. equal to the number of coin tosses we need to get the first heads. Let A be the event that we get a heads in the first toss. Using the law of total expectation,

$$\mathbb{E}[R^2] = \mathbb{E}[R^2|A] \Pr[A] + \mathbb{E}[R^2|A^c] \Pr[A^c]$$

$\mathbb{E}[R^2|A] = 1$ ($R^2 = 1$ if we get a heads in the first coin toss) and $\Pr[A] = p$. Hence,

$$\mathbb{E}[R^2] = (1)\,(p) + \mathbb{E}[R^2|A^c]\,(1-p) \quad ; \quad \mathbb{E}[R^2|A^c] = \sum_{k=1} k^2 \Pr[R = k|A^c]$$

Note that $\Pr[R = k|A^c] = \Pr[R = k| \text{ first toss is a tails}] = (1-p)^{k-2}\, p = \Pr[R = k-1]$

$$\implies \mathbb{E}[R^2|A^c] = \sum_{k=1} k^2 \Pr[R = k-1] = \sum_{t=0} (t+1)^2 \Pr[R = t] \qquad (t := k-1)$$

Variance - Examples

Continuing from the previous slide,

$$\mathbb{E}[R^2|A^c] = \sum_{t=0}(t+1)^2 \Pr[R=t] = \sum_{t=0} t^2 \Pr[R=t] + 2\sum_{t=0} t \Pr[R=t] + \sum_{t=0} \Pr[R=t]$$

$$= \sum_{t=1} t^2 \Pr[R=t] + 2\sum_{t=1} t \Pr[R=t] + \sum_{t=1} \Pr[R=t] = \mathbb{E}[R^2] + 2\mathbb{E}[R] + 1$$

Putting everything together,

$$\mathbb{E}[R^2] = (1)\,(p) + (\mathbb{E}[R^2] + 2\mathbb{E}[R] + 1)\,(1-p) \implies p\,\mathbb{E}[R^2] = p + 2(1-p)\mathbb{E}[R] + (1-p)$$

$$\implies p\,\mathbb{E}[R^2] = p + \frac{2(1-p)}{p} + (1-p) \qquad (\mathbb{E}[R] = \tfrac{1}{p})$$

$$\implies \mathbb{E}[R^2] = \frac{2(1-p)}{p^2} + \frac{1}{p} \implies \mathbb{E}[R^2] = \frac{2-p}{p^2}$$

$$\implies \mathrm{Var}[R] = \mathbb{E}[R^2] - (\mathbb{E}[R])^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Properties of Variance

Q: For constants a, b and r.v. R, prove that $\mathrm{Var}[a\,R + b] = a^2\mathrm{Var}[R]$.

Proof:

$$\begin{aligned}\mathrm{Var}[aR+b] &= \mathbb{E}[(aR+b)^2] - (\mathbb{E}[aR+b])^2 = \mathbb{E}[a^2R^2 + 2abR + b^2] - (\mathbb{E}[aR] + \mathbb{E}[b])^2 \\ &= (a^2\mathbb{E}[R^2] + 2ab\mathbb{E}[R] + b^2) - (a\mathbb{E}[R] + b)^2 \\ &= (a^2\mathbb{E}[R^2] + 2ab\mathbb{E}[R] + b^2) - (a^2(\mathbb{E}[R])^2 + 2ab\mathbb{E}[R] + b^2) \\ &= a^2\left[\mathbb{E}[R^2] - (\mathbb{E}[R])^2\right] \\ \implies \mathrm{Var}[aR+b] &= a^2\mathrm{Var}[R]\end{aligned}$$

- Similarly, for the standard deviation,

$$\sigma_{aR+b} = \sqrt{\mathrm{Var}[aR+b]} = \sqrt{a^2\mathrm{Var}[R]} = |a|\,\sigma_R$$

- Note the difference from the property of expectation,

$$\mathbb{E}[aR+b] = a\mathbb{E}[R] + b$$

Properties of Variance

Q: If the r.v's R_1 and R_2 are *independent*, prove that $\text{Var}[R_1 + R_2] = \text{Var}[R_1] + \text{Var}[R_2]$.

- In order to prove this result, we need an additional definition: for 2 r.v's R_1 and R_2,

$$\mathbb{E}[R_1\, R_2] := \sum_{z \in \text{Range}(R_1\, R_2)} z \, \Pr[R_1\, R_2 = z] = \sum_{x \in \text{Range}(R_1)} \sum_{y \in \text{Range}(R_2)} x\, y \, \Pr[R_1 = x \cap R_2 = y]$$

Proof: We will first prove that for two independent r.v's R_1 and R_2, $\mathbb{E}[R_1\, R_2] = \mathbb{E}[R_1]\,\mathbb{E}[R_2]$.

$$\begin{aligned}
\mathbb{E}[R_1 R_2] &= \sum_{x \in \text{Range}(R_1)} \sum_{y \in \text{Range}(R_2)} x\, y \, \Pr[R_1 = x \cap R_2 = y] \\
&= \sum_{x \in \text{Range}(R_1)} \sum_{y \in \text{Range}(R_2)} xy \, \Pr[R_1 = x] \Pr[R_2 = y] && \text{(Independence)} \\
&= \sum_{x \in \text{Range}(R_1)} x \Pr[R_1 = x] \sum_{y \in \text{Range}(R_2)} y \Pr[R_2 = y] \\
&= \mathbb{E}[R_1]\, \mathbb{E}[R_2]
\end{aligned}$$

Properties of Variance

Continuing the proof from the previous slide and using the definition of $\text{Var}[R_1 + R_2]$,

$$
\begin{aligned}
\text{Var}[R_1 + R_2] &= \mathbb{E}[(R_1 + R_2)^2] - (\mathbb{E}[R_1 + R_2])^2 \\
&= \mathbb{E}[R_1^2 + R_2^2 + 2R_1\, R_2] - (\mathbb{E}[R_1] + \mathbb{E}[R_2])^2 \\
&= \mathbb{E}[R_1^2 + R_2^2 + 2R_1\, R_2] - [(\mathbb{E}[R_1])^2 + (\mathbb{E}[R_2])^2 + 2\mathbb{E}[R_1]\,\mathbb{E}[R_2]] \\
& \qquad \text{(Expanding the terms)} \\
&= [\mathbb{E}[R_1^2] - (\mathbb{E}[R_1])^2] + [\mathbb{E}[R_2^2] - (\mathbb{E}[R_2])^2] + 2(\mathbb{E}[R_1\, R_2] - \mathbb{E}[R_1]\,\mathbb{E}[R_2]) \\
& \qquad \text{(Linearity of expectation)} \\
&= \text{Var}[R_1] + \text{Var}[R_2] + 2(\mathbb{E}[R_1\, R_2] - \mathbb{E}[R_1]\,\mathbb{E}[R_2]) \quad \text{(Definition of variance)}
\end{aligned}
$$

From the previous slide, if R_1 and R_2 are independent, then, $\mathbb{E}[R_1\, R_2] = \mathbb{E}[R_1]\,\mathbb{E}[R_2]$,

$$\implies \text{Var}[R_1 + R_2] = \text{Var}[R_1] + \text{Var}[R_2]$$

Properties of Variance

Q: For pairwise independent random variables $R_1, R_2, R_3, \dots R_n$, $\text{Var}[\sum_{i=1}^n R_i] = \sum_{i=1}^n \text{Var}[R_i]$.

Proof: Following the same proof, we can show that for any pair of pairwise independent r.v's, R_i and R_j, $\mathbb{E}[R_i R_j] = \mathbb{E}[R_i]\,\mathbb{E}[R_j]$.

$$\text{Var}[R_1 + R_2 + \dots R_n] = \mathbb{E}[(R_1 + R_2 + \dots R_n)^2] - (\mathbb{E}[R_1 + R_2 + \dots R_n])^2$$

$$= \sum_{i=1}^n [\mathbb{E}[R_i^2] - (\mathbb{E}[R_i])^2] + 2 \sum_{i,j|1 \le i < j \le n} [\mathbb{E}[R_i R_j] - \mathbb{E}[R_i]\,\mathbb{E}[R_j]]$$

(Expanding the terms and using linearity of expectation)

$$\implies \text{Var}[R_1 + R_2 + \dots R_n] = \sum_{i=1}^n \text{Var}[R_i]$$

(Since the r.v's are pairwise independent)

- In general, the pairwise independence of r.v.'s is a necessary condition for the linearity of variance. To see this, consider $R_1 = R_2 = R$ i.e. the two r.v's are not independent. In this case, $\text{Var}[R_1 + R_2] = \text{Var}[2R] = 4\text{Var}[R] \neq 2\text{Var}[R] = \text{Var}[R_1] + \text{Var}[R_2]$.

Variance - Examples

Q: If $R \sim \text{Bin}(n, p)$, calculate $\text{Var}[R]$.

Define R_i to be the indicator random variable that we get a heads in toss i of the coin. Recall that R is the random variable equal to the number of heads in n tosses.

Hence,

$$R = R_1 + R_2 + \ldots + R_n \implies \text{Var}[R] = \text{Var}[R_1 + R_2 + \ldots + R_n]$$

Since $R_1, R_2, \ldots, R_n$ are mutually independent indicator random variables and mutual independence implies pairwise independence,

$$\text{Var}[R] = \text{Var}[R_1] + \text{Var}[R_2] + \ldots + \text{Var}[R_n]$$

Since the variance of an indicator (Bernoulli) r.v. is $p(1-p)$,

$$\text{Var}[R] = n\,p\,(1-p).$$

Questions?