CS70: Jean Walrand: Lecture 28.

Independence and Variance

1. Expected value of nonnegative integer-valued RV
2. Waiting for a bus
3. A note about expectation
4. Independent random variables.
5. Variance.

Summary of Lecture 27

**Expected Value of Nonnegative Integer RV**

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} \text{Pr}[X \geq i].$$

**Proof:** One has

$$E[X] = \sum_{i=1}^{\infty} i \times \text{Pr}[X = i]$$

Using (1) gives

$$E[X] = \sum_{i=1}^{\infty} i \times \text{Pr}[X = i] - \sum_{i=1}^{\infty} (i - 1) \times \text{Pr}[X = i] = \sum_{i=1}^{\infty} \text{Pr}[X \geq i].$$

A note about expectation

Recall that

$$E[X] = \sum_{x} x \times \text{Pr}[X = x] = \sum_{\omega} X(\omega) \times \text{Pr}[\omega]. \quad (1)$$

**Theorem**

(a) $E[g(X)] = \sum_{x} g(x) \times \text{Pr}[X = x].$

(b) $E[g(X, Y, Z)] = \sum_{x, y, z} g(x, y, z) \times \text{Pr}[X = x, Y = y, Z = z].$

**Proof** (a) One has

$$E[g(X)] = \sum_{x} g(X(\omega)) \times \text{Pr}[\omega] \text{ by (1) applied to } Y(\omega) = g(X(\omega))$$

$$= \sum_{x} \sum_{\omega: X(\omega) = x} g(X(\omega)) \times \text{Pr}[\omega] = \sum_{x} \sum_{\omega: X(\omega) = x} g(x) \times \text{Pr}[\omega] = \sum_{x} g(x) \times \text{Pr}[X = x].$$

(b) is similar.

Riding the bus.

How about the waiting time until the next bus?

Add the black dot uniformly at random and pretend that it represents your arrival time. This is legitimate, because given the black dot, the other dots are uniform at random. Then,

$$24 = E[X_1 + \cdots + X_n] = n \times E[X_1],$$

by linearity and symmetry.

Hence, $E[X_1] = E[X] = 24/5 = 24/(n+1)$ for $n$ busses. Note the paradox: the average time between the previous and next bus is $48/(n+1) > 24/n$.

Riding the bus.

$n$ buses arrive uniformly at random throughout a 24 hour day. What is the time between buses? How long do you expect to wait for a bus? Here are typical arrival times, independent and uniform in $[0, 24]$.

Here is an alternative picture (left)

The $n$ gaps between busses add up to 24. Thus $n$ times the expected gap is 24 and the expected gap between busses is $24/n$. 

Expected Value of Nonnegative Integer RV

**Theorem:** For a r.v. $X$ that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} \text{Pr}[X \geq i].$$
Independent Random Variables.

**Definition:** Independence

The random variables $X$ and $Y$ are independent if and only if

$$\Pr[Y = b|X = a] = \Pr[Y = b], \text{ for all } a \text{ and } b.$$  

**Fact:**

$X, Y$ are independent if and only if

$$\Pr[X = a, Y = b] = \Pr[X = a]\Pr[Y = b], \text{ for all } a \text{ and } b.$$  

Indeed:

$$\Pr[Y = b|X = a] = \frac{\Pr[X = a, Y = b]}{\Pr[X = a]} = \Pr[Y = b]$$

$\implies \Pr[X = a, Y = b] = \Pr[X = a]\Pr[Y = b].$

---

### Independence: Examples

**Example 1**

Roll two dices. $X, Y =$ number of pips on the two dice. $X, Y$ are independent.

Indeed: $\Pr[X = a, Y = b] = \Pr[X = a, Y = b] = \Pr[X = a]\Pr[Y = b].$

**Example 2**

Roll two dices. $X =$ total number of pips, $Y =$ number of pips on die 1 minus number on die 2. $X$ and $Y$ are not independent.

Indeed: $\Pr[X = 12, Y = 1] = 0 \neq \Pr[X = 12]\Pr[Y = 1].$

**Example 3**

Flip a fair coin five times, $X =$ number of Hs in first three flips, $Y =$ number of Hs in last two flips. $X$ and $Y$ are independent.

Indeed:

$$\Pr[X = a, Y = b] = \binom{3}{a}\binom{2}{b}2^{-5} = \binom{3}{a}\binom{2}{b}2^{-5} = \Pr[X = a]\Pr[Y = b].$$

---

### Functions of Independent Random Variables

**Theorem**

Functions of independent RVs are independent.

Let $X, Y$ be independent RV. Then $f(X)$ and $g(Y)$ are independent, for all $f(\cdot), g(\cdot)$.

**Proof:**

Recall the definition of inverse image:

$$h(z) \in \mathbb{C} \iff z \in h^{-1}(C) := \{ z \mid h(z) \in C \}. \quad (2)$$

Now,

$$\Pr[f(X) \in A, g(Y) \in B] = \Pr[X \in f^{-1}(A), Y \in g^{-1}(B)], \text{ by (2)}$$

= $\Pr[X \in f^{-1}(A)]\Pr[Y \in g^{-1}(B)],$ since $X, Y$ ind.

= $\Pr[f(X) \in A]\Pr[g(Y) \in B], \text{ by (2)}.$

---

### Mean of product of independent RV

**Theorem**

Let $X, Y$ be independent RVs. Then

$$E[XY] = E[X]E[Y].$$

**Proof:**

Recall that $E[g(X, Y)] = \sum_{x,y} g(x, y)\Pr[X = x, Y = y]$. Hence,

$$E[XY] = \sum_{x,y} xy\Pr[X = x, Y = y] = \sum_{x} \sum_{y} xy\Pr[X = x]\Pr[Y = y], \text{ by ind.}$$

= $\sum_{x} \sum_{y} \Pr[X = x]\Pr[Y = y] = \sum_{x} \Pr[X = x]\sum_{y} \Pr[Y = y] = \sum_{x} \Pr[X = x]E[Y] = E[X]E[Y].$

---

### A useful observation about independence

**Theorem**

$X$ and $Y$ are independent if and only if

$$\Pr[X \in A, Y \in B] = \Pr[X \in A]\Pr[Y \in B] \text{ for all } A, B \subset \mathbb{R}. \quad (1)$$

**Proof:**

If $\iff$ Choose $A = \{a\}$ and $B = \{b\}$.

This shows that $\Pr[X = a, Y = b] = \Pr[X = a]\Pr[Y = b].$

Only if $\iff$:

$$\Pr[X \in A, Y \in B] = \sum_{a \in A} \sum_{b \in B} \Pr[X = a, Y = b] = \sum_{a \in A} \sum_{b \in B} \Pr[X = a]\Pr[Y = b] = \sum_{a \in A} \Pr[X = a]\Pr[Y \in B] = \Pr[X \in A]\Pr[Y \in B].$$

**Examples**


Then

$$E[(X + 2Y + 3Z)^2] = E[X^2] + 4Y^2 + 9Z^2 + 4XY + 12YZ + 6XZ]$$

= $1 + 4 + 9 + 4 \times 0 + 12 \times 0 + 6 \times 0$

= $14.$

(2) Let $X, Y$ be independent and $U[1, 2, \ldots, n]$. Then


= $\frac{1 + 3n + 2n^2}{3} - \frac{(n+1)^2}{2}.$
### Mutually Independent Random Variables

**Definition**

(a) $X, Y, Z$ are mutually independent if
\[
\Pr[X = x, Y = y, Z = z] = \Pr[X = x]\Pr[Y = y]\Pr[Z = z], \quad \text{for all } x, y, z.
\]

(b) $\{X_j, j \in J\}$ are mutually dependent if
\[
\Pr[X_k = x_k, k \in K] = \prod_{k \in K}\Pr[X_k = x_k], \quad \text{for all finite } K \subset J \text{ and all } x_k.
\]

Note: If $A, B, C$ are pairwise independent but not mutually independent, then $X = 1_A, Y = 1_B, Z = 1_C$ are pairwise independent but not mutually independent.

### Variance and Standard Deviation

**Fact:**
\[
\]

Indeed:
\[
\begin{align*}
\text{var}(X) &= E[(X - E[X])^2] \\
&= E[X^2 - 2XE[X] + E[X]^2] \\
&= E[X^2] - 2E[X]E[X] + E[X]^2, \quad \text{by linearity} \\
&= E[X^2] - E[X]^2.
\end{align*}
\]

Consider the random variable $X$ such that
\[
X = \begin{cases} 
\mu - \sigma, & \text{w.p. } 1/2 \\
\mu + \sigma, & \text{w.p. } 1/2.
\end{cases}
\]

Then, $E[X] = \mu$ and $(X - E[X])^2 = \sigma^2$. Hence,
\[
\text{var}(X) = \sigma^2 \quad \text{and} \quad \sigma(X) = \sigma.
\]

### Functions of mutually independent RVs

If $X, Y, Z$ are pairwise independent, but not mutually independent, it may be that $f(X)$ and $g(Y, Z)$ are not independent.

**Example:** Flip two fair coins, $X = 1$ (coin 1 is $H$), $Y = 1$ (coin 2 is $H$), $Z = X \oplus Y$. Then, $X, Y, Z$ are pairwise independent. Let $g(Y, Z) = Y \oplus Z$. Then $g(Y, Z) = X$ is not independent of $X$.

**Theorem**

Let $\{X_j, j \in J\}$ be mutually independent RVs. Then
\[
t_1(X_1), t_2(X_2, j \in K_2) \ldots \text{ are mutually independent}
\]

for all pairwise disjoint $K_m$ and functions $t_m(\cdot)$.

**Proof:**

Try it, you will like it. \(\square\)

### Variance

The variance measures the deviation from the mean value.

**Definition:** The variance of $X$ is
\[
\sigma^2(X) := \text{var}[X] = E[(X - E[X])^2].
\]

$\sigma(X)$ is called the standard deviation of $X$.

### A simple example

This example illustrates the term 'standard deviation.'

```
\begin{align*}
\Pr &\quad 0.5 \\
\mu - \sigma &\quad \sigma &\quad \mu + \sigma \\
\end{align*}
```

Consider the random variable $X$ such that
\[
X = \begin{cases} 
\mu - \sigma, & \text{w.p. } 1/2 \\
\mu + \sigma, & \text{w.p. } 1/2.
\end{cases}
\]

Then, $E[X] = \mu$ and $(X - E[X])^2 = \sigma^2$. Hence,
\[
\text{var}(X) = \sigma^2 \quad \text{and} \quad \sigma(X) = \sigma.
\]

### Example

Consider $X$ with
\[
X = \begin{cases} 
-1, & \text{w. p. } 0.99 \\
99, & \text{w. p. } 0.01.
\end{cases}
\]

Then
\[
E[X] = -1 \times 0.99 + 99 \times 0.01 = 0.
\]
\[
E[X^2] = 1 \times 0.99 + (99)^2 \times 0.01 \approx 100.
\]
\[
\text{Var}(X) \approx 100 \implies \sigma(X) \approx 10.
\]

Also,
\[
E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.
\]

Thus, $\sigma(X) \neq E[|X - E[X]|]$.

**Exercise:** How big can you make $\sigma(X)$?
Uniform
Assume that $Pr[X = i] = 1/n$ for $i \in \{1, \ldots, n\}$. Then

$$E[X] = \frac{\sum i \times Pr[X = i]}{n} = \frac{1}{n} \sum i = \frac{n}{2}$$

$$E[X^2] = \frac{\sum i^2 \times Pr[X = i]}{n} = \frac{1}{n} \sum i^2 = \frac{1 + 3n + 2n^2}{6}$$

This gives

$$var(X) = \frac{1 + 3n + 2n^2}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

**Properties of variance.**

1. $Var(cX) = c^2 Var(X)$, where $c$ is a constant.
   Scales by $c^2$.
2. $Var(X + c) = Var(X)$, where $c$ is a constant.
   Shifts center.

**Proof:**

\[
Var(cX) = E((cX)^2) - (E(cX))^2 = c^2 E(X^2) - c^2 (E(X))^2 = c^2 Var(X)
\]

\[
Var(X + c) = E((X + c)^2) - (E(X + c))^2 = E((X + c)^2) - (E(X) + c)^2 = E((X - E(X))^2) = Var(X)
\]

Fixed points.
Number of fixed points in a random permutation of $n$ items.
"Number of student that get homework back."

$X = X_1 + X_2 \cdots + X_n$

where $X_i$ is indicator variable for $i$th student getting hw back.

$$E(X^2) = \sum_i E(X_i^2) + \sum_{ij} E(X_iX_j) = n \times \frac{1}{n} + (n-1) \times \frac{1}{n(n-1)} = 1 + 1 = 2.$$ 

$$Var(X) = E(X^2) - (E(X))^2 = 2 - 1 = 1.$$ 

Variance of sum of independent random variables

**Theorem:**
If $X$ and $Y$ are independent, then

$$Var(X + Y) = Var(X) + Var(Y).$$

**Proof:**
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X) = 0$ and $E(Y) = 0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$ 

Hence,

$$Var(X + Y) = Var((X + Y)^2) = Var(X^2 + 2XY + Y^2) = Var(X^2) + 2E(XY) + Var(Y^2) = Var(X) + Var(Y).$$

Variance: binomial.

$$E[X^2] = \sum_{i=0}^n i^2 \left(\frac{n!}{(n-i)!i!}\right) p^i (1-p)^{n-i}.$$ 

Too hard!
Ok.. fine.
Let's do something else.
Maybe not much easier...but there is a payoff.

Variance of Binomial Distribution.

Flip coin with heads probability $p$.

$X$- how many heads?

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X_i^2] = 1^2 \times p + 0^2 \times (1-p) = p.$$ 

$$Var(X_i) = p \times (E[X_i])^2 = p - p^2 = p(1-p).$$ 

$p = 0 \implies Var(X_i) = 0$

$p = 1 \implies Var(X_i) = 0$

$X = X_1 + X_2 + \ldots + X_n$.

$X_i$ and $X_j$ are independent: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1]$.

$$Var(X) = Var(X_1 + \cdots + X_n) = np(1-p).$$
Summary

**Independence and Variance**

- \( E[g(X, Y, Z)] = \sum_{x,y,z} g(x, y, z) Pr[X = x, Y = y, Z = z] \)
- \( X, Y \) independent \( \iff \) \( Pr[X \in A, Y \in B] = Pr[X \in A] Pr[Y \in B] \)
- Then, \( f(X), g(Y) \) are independent
- Also, \( E[XY] = E[X]E[Y] \) and \( \text{var}[X + Y] = \text{var}[X] + \text{var}[Y] \).

**Distributions:**

- \( U[1, \ldots, n] : E[X] = (n + 1)/2; \text{var}[X] = (n^2 - 1)/12 \)
- \( B(n, p) : E[X] = np; \text{var}[X] = np(1 - p) \).