Algorithms and Probability (FS2025)
Week 7

Rui Zhang

April 2, 2025

Contents

1 Peer Grading Feedback

2 Content

2.1 Conditional- / Multiple- / Independent- Random Variables

Definition: Let X be a random variable on Ω, $A \subseteq \Omega$ an event such that $Pr[A] > 0$. $X \mid A$ is what we call a conditional random variable. if $X : \Omega \to \mathbb{R}$, then we have that

$$X \mid A : A \to \mathbb{R}$$

So, the domain of the function X has been reduced to A.

Now that we have this new form of random variables, we have to also make sure that the distribution, distribution function and the expectation (all concepts we have learnt until now) are still defined:

Definition: We will now define the same concepts as before, this time for conditional random variables:

- $f_{X|A} = \Pr[X = x|A]$
- $F_{X|A} = \Pr[X \leq x|A]$
- $E[X|A] = \sum_{x \in W_X} x \cdot \Pr[X = x|A] = \frac{1}{\Pr[A]} \sum_{\omega \in A} X(\omega) \cdot \Pr[\omega]$

And we say that X is independent of A, if $f_{X|A} = f_X$

Definition:

- $f_{X,Y}(x, y) := \Pr[X = x, Y = y]$
- $f_X(x) := \sum_{y \in W_Y} f_{X,Y}(x, y)$

Exercise: Adapted from FS2021:

Let $n \geq 2$. Consider the following random process: In the first step, X is drawn uniformly at random from the set $\{1, 2, \ldots, n\}$. In the second step, Y is drawn uniformly at random from the set $\{1, \ldots, X\}$.

(a) How many possible values exist for the pair (X, Y)?

$X \in \{1, \ldots, n\}$ and $Y \in \{1, \ldots, X\}$ Thus $\sum_{x=1}^{n} x = \frac{n(n+1)}{2}$ total options

(b) Determine the joint probability mass function of X and Y.

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{n} \cdot \frac{1}{x}, & 1 \leq y \leq x \leq n \\ 0 & \text{else} \end{cases}$$

$\frac{1}{n} \cdot \frac{1}{x} = \Pr[X = x] \cdot \Pr[Y = y \mid X = x]$

$= \Pr[X = x, Y = y]$

(c) Let $y \in \{1, \ldots, n\}$. Compute $\Pr[Y = y]$.

$$\Pr[Y = y] = f_Y(y) = \sum_{x \in W_X} f_{X,Y}(x,y) = \sum_{y \leq x \leq n} f_{X,Y}(x,y) = \sum_{y \leq x \leq n} \frac{1}{n} \cdot \frac{1}{x} = \frac{1}{n} \sum_{y \leq x \leq n} \frac{1}{x} = \frac{1}{n}\left(\sum_{x=1}^{n} \frac{1}{x} - \sum_{x=1}^{y-1} \frac{1}{x}\right)$$

$$= \frac{1}{n}(H_n - H_{y-1})$$

Independence:

Definition: Let $X_1, \ldots, X_n$ be random variables. They are called independent if and only if

$$\Pr[X_1 = x_1, \ldots, X_n = x_n] = \Pr[X_1 = x_1] \cdots \Pr[X_n = x_n]$$

for all $x_i \in W_{X_i}$.

For independent random variables, we have

> **Lemma:** Let $X_1, \ldots, X_n$ be independent random variables, then it holds that:
>
> $$\Pr[X_1 \in S_1, \ldots, X_n \in S_n] = \Pr[X_1 \in S_1] \cdots \Pr[X_n \in S_n]$$
>
> for all subsets $S_i \subseteq \mathbb{R}$

Note that this is not a definition but a characteristic of random variables that we have already established to be independent. In particular, we can derive that two Indicator Variables $\mathbb{I}_A$ and $\mathbb{I}_B$ are independent when the following holds:

> **Lemma:** $\mathbb{I}_A$ and $\mathbb{I}_B$ are independent if and only if
>
> $$f_{X,Y}(1,1) = f_X(1) \cdot f_Y(1)$$
>
> And we do not have to check the equation for $x = 0$ or $y = 0$!

Independence of random variables is furthermore preserved under function application!

> **Theorem:** Let $X_1, \ldots, X_n$ be independent random variables and $f_1, \ldots f_n : \mathbb{R} \to \mathbb{R}$ real functions, then $f_1(X_1), \ldots f_n(X_n)$ are also independent random variables.

Let us continue the previous exercise:

(d) Determine the conditional probability $\Pr[Y = y \mid X = x]$ for all $x, y \in \{1, \ldots, n\}$.

$$\mathbb{P}[Y=y \mid X=x] = \begin{cases} \frac{1}{x}, & y \leq x \\ 0, & y > x \end{cases} \quad \text{(from before)}$$

(e) Determine the conditional probability $\Pr[X = x \mid Y = y]$ for all $x, y \in \{1, \ldots, n\}$.

$$\mathbb{P}[X=x \mid Y=y] = \frac{\mathbb{P}[X=x \cap Y=y]}{\mathbb{P}[Y=y]} = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{\cancel{n}x} \cdot \cancel{n} \frac{1}{H_n - H_{y-1}} & 1 \leq y \leq x \leq n \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} \frac{1}{x(H_n - H_{y-1})} & 1 \leq y \leq x \leq n \\ 0 & \text{else} \end{cases}$$

(f) Show that X and Y are not independent.

$$\mathbb{P}[X=n, Y=n] \overset{!}{=} \mathbb{P}[X=n]\mathbb{P}[Y=n],$$

But $\frac{1}{n^2} \neq \frac{1}{n} \cdot \frac{1}{n^2}$

> **Theorem:** Let X and Y be two independent random variables and let $Z := X + Y$, then we have that
>
> $$f_Z(z) = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)$$

Theorem: (Wald's Equation) Let N and X be two independent random variables, where the range of N is the natural numbers and $Z := \sum_{i=1}^{N} X_i$ and $X_1, X_2, \ldots$ are independent copies of X. Then we have that

$$\mathbb{E}[Z] = \mathbb{E}[N]\mathbb{E}[X]$$

Proof. □

$$\mathbb{E}[Z] = \sum_{n \in \mathbb{N}} \mathbb{E}[Z \mid N = n]\, \mathbb{P}[N = n] = \sum_{n \in \mathbb{N}} \mathbb{E}\Big[\sum_{i=1}^{n} X_i\Big]\, \mathbb{P}[N = n]$$

$$= \sum_{n \in \mathbb{N}} \sum_{i=1}^{n} \mathbb{E}[X_i]\, \mathbb{P}[N = n]$$

$$= \sum_{n \in \mathbb{N}} \sum_{i=1}^{n} \mathbb{E}[X]\, \mathbb{P}[N = n]$$

$$= \mathbb{E}[X] \sum_{n \in \mathbb{N}} n \cdot \mathbb{P}[N = n]$$

$$= \mathbb{E}[X]\,\mathbb{E}[N]$$

2.2 Distributions

Mathematicians like to produce general very statements about what they create. Thus we introduce the notion of distributions. A random variable can follow a certain distribution and if we notice that, then we suddenly know a lot more about the nature of this random variable.

Definition: A random variable X is said to be a Bernoulli-Variable with parameter $0 \leq p \leq 1$ (Bernoulli-verteilt mit Parameter p, $X \sim \text{Bernoulli}(p)$), if and only if we have

$$f_X(x) = \begin{cases} p \text{ if } x = 1 \\ 1-p \text{ if } x = 0 \\ 0 \text{ else} \end{cases}$$

In which case $\mathbb{E}[X] = p$.

Definition: A random variable X is said to follow a binomial distribution (Binomialverteilt, $X \sim \text{Bin}(n,p)$) if and only if

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, \; x \in \{0, 1, \dots, n\} \\ 0, \text{else} \end{cases}$$

And we have $\mathbb{E}[X] = n \cdot p$

Definition: A random variable X is said to follow a poisson distribution (Poisson-Verteilt, $X \sim \text{Poi}(\lambda)$) if and only if

$$f_X(i) = \begin{cases} \frac{e^{-\lambda}\lambda^i}{i!} \text{ for } i \in \mathbb{N}_0 \\ 0 \text{ else} \end{cases}$$

Then $\mathbb{E}[X] = \lambda$. We have that $\lim_{n\to\infty} \text{Bin}(n, \lambda/n) = \text{Poi}(\lambda)$

Definition: A random variable X is said to be a geometric random variable with parameter p ($X \sim \text{Geo}(p)$) if and only if

$$f_X(i) = \begin{cases} (1-p)^{i-1} p \text{ for } i \in \mathbb{N} \\ 0 \text{ else} \end{cases}$$

And then we have $\mathbb{E}[X] = 1/p$.

Theorem: $X \sim \text{Geo}(p)$, then we have $\forall s, t \in \mathbb{N}$:

$$\Pr[X \geq s + t \mid X > s] = \Pr[X \geq t]$$

Definition: A random variable X is said to be a negative binomial random variable with parameter n and p ($X \sim \text{NegativeBinomial}(n,p)$) if and only if

$$f_X(k) = \begin{cases} \binom{k-1}{n-1} (1-p)^{k-n} p^n \text{ for k=1,2,...} \\ 0 \text{ else} \end{cases}$$

and then $\mathbb{E}[X] = n/p$

Short Quizzes & Mini Quiz Prep (Together)

- Let $X_1 \sim \text{Bin}(n, p)$ and $X_2 \sim \text{Bin}(n, p)$. Which of the following statements is true?
  - $X_1 + X_2 \sim \text{Bin}(n, 2p)$
  - $X_1 + X_2 \sim \text{Bin}(2n, p)$ ✓ $\Rightarrow$ follows from previous theorem:
  - $X_1 + X_2 \sim \text{Bin}(n, p)$
  - $X_1 + X_2$ does not necessarily have to follow a binomial distribution
  - $X_1 + X_2 \sim \text{Bin}(2n, 2p)$

Correction: this is correct, bc. X_1 & X_2 have to be independent...

$$Z = X_1 + X_2$$

$$f_Z(a) = \sum_{b=0}^{n} f_{X_1}(b) \cdot f_{X_2}(a-b)$$
$$= \sum_{b=0}^{n} \binom{n}{b} p^b (1-p)^{n-b} \binom{n}{a-b} p^{a-b} (1-p)^{n-a+b}$$
$$= \sum_{b=0}^{n} \binom{n}{b} \binom{2n-n}{a-b} p^a (1-p)^{2n-a}$$
$$= p^a (1-p)^{2n-a} \sum_{b=0}^{n} \binom{n}{b} \binom{2n-n}{a-b}$$
$$= \binom{2n}{a} p^a (1-p)^{2n-a}$$

- Every Indicator Variable is a Bernoulli-Variable ✓
- We roll a dice 4 times in a row. The sum of results follows a binomial distribution ✗

2.3 Variance

Definition: Let X be a random variable with expectation $\mu = \mathbb{E}[X]$. We then define the variance of X to be

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

and we call the square root of this value the standard deviation $\sigma = \sqrt{\text{Var}[X]}$

Theorem: $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Theorem: $\text{Var}[a \cdot X + b] = a^2 \cdot \text{Var}[X]$

2.4 Exercise: Deriving Target Shooting Together!

With the knowledge of distributions we have now, we can derive Expectation and Variance of the "Target Shooting" Problem, which you guys will hopefully see soon!

Given a set U and a subset $S \subseteq U$ of unknown size, how large is $|S|/|U|$? Target Shooting aims to determine this information by randomly choosing elements u from U, and determining whether $u \in S$. If you repeat this procedure often enough, you will be able to get an estimate of the ratio $|S|/|U|$. Formally, Target Shooting uses the following algorithm:

Algorithm 1 Target-Shooting

```
Choose u_1, ..., u_N ∈ U randomly, uniformly distributed, and independently
return (1/N) · Σ_{i=1}^{N} I_S(u_i)
```

Show that the expectation of the returned value from this algorithm is $|S|/|U|$ and derive an equation for the Variance:

Expectation:

write in class

Let Z := "Value returned by Algo"

$Z = \frac{1}{N} \sum_{i=1}^{N} I_S(u_i)$.

What are the $I_S(u_i)$'s and what is their distribution?

What is $\sum_{i=1}^{N} I_S(u_i)$ and what is its distribution?

What is $E[Z]$?

$\Rightarrow E[Z] = \frac{1}{N} E[Z']$, $Z' \sim Bin(N, \frac{|S|}{|U|})$

$= \frac{1}{\cancel{N}} \cancel{N} \cdot \frac{|S|}{|U|} = \frac{|S|}{|U|}$ yey.

Variance: ~~("Double Sum" Trick)~~

$$Var[Z] = \frac{1}{N^2} Var[Z']$$

$$= \frac{1}{N^{\cancel{2}}} \cdot \left(\cancel{N} \frac{|S|}{|U|} \left(1 - \frac{|S|}{|U|} \right) \right)$$

$$= \frac{1}{N} \left(\frac{|S|}{|U|} - \left(\frac{|S|}{|U|} \right)^2 \right)$$ b.c. of binomial dist.

	$i=1$	$i=2$	$i=3$	$\dots$	i=n
$\frac{1}{n-i+1}$	$1/n$	$1/(n-1)$	$1/(n-2)$	$\dots$	1
$\frac{1}{i}$	1	$1/2$	$1/3$	$\dots$	$1/n$

Figure 1: Visualisation for what is happening in step $(*)$ for the derivation of coupon collector. What you'll realize is that we actually still sum up over the same numbers.

2.5 Coupon Collector

Consider n cards we want to collect. We want to calculate the number of cards X we have to buy in order to collect all of the cards. We will write X as a sum of random variables:

$$X = X_1 + X_2 + \dots X_n$$

$$X_1 = 1$$
$$X_2 \sim Geo\left(\frac{n-1}{n}\right)$$
$$X_3 \sim Geo\left(\frac{n-2}{n}\right)$$
$$\vdots$$
$$X_n \sim Geo\left(\frac{1}{n}\right)$$

$$X_i \sim \text{Geo}\left(\frac{n-(i-1)}{n}\right)$$

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

due to linearity of expectation. And we can continue calculating and get:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] \\
&= \sum_{i=1}^{n} \frac{n}{n-i+1} \\
&= n \sum_{i=1}^{n} \frac{1}{n-i+1} \quad \Leftarrow \text{start here!} \\
&\overset{(*)}{=} n \sum_{i=1}^{n} \frac{1}{i} \\
&= n \cdot H_n \\
&= n \cdot \ln n + O(n)
\end{aligned}$$

where H_n is roughly $\ln n + O(1)$.

Short Quizzes & Mini Quiz Prep (3min, choose one)

- Take the coupon collector problem with $n = 3$ coupons. Calculate the expected number of rounds to get all coupons.

$$3\cdot\sum_{i=1}^{3}\frac{1}{i} = 3\left(1+\frac{1}{2}+\frac{1}{3}\right) = 3+1.5+1 = 5.5$$

- Rewe has released new cards again! There are a total of n cards and you were lucky enough to be gifted $\frac{n}{2}$ different coupons. The expected number of needed rounds can be expressed as $c \cdot nlnn + O(n)$ for some constant c. Determine c.

$$n\sum_{i=\frac{n}{2}}^{n}\frac{1}{n-i+1} = n\sum_{i=1}^{\frac{n}{2}}\frac{1}{i} = n\cdot H_{\frac{n}{2}} = n\ln\left(\frac{n}{2}\right) + O(n) = n\left(\ln(n) - \ln(2)\right) + O(n)$$

$$= n\ln n - n\ln(2) + O(n) \Rightarrow c = 1$$

- Aldi has released new cards again! There are a total of n cards and you were lucky enough to get a voucher which you can redeem when you have 90% of the cards to get the entire rest of the cards. The number of rounds needed to get all the rounds using the voucher can be expressed as $c \cdot nlnn + O(n)$ for some constant c. Determine c.

$$n\sum_{i=1}^{\frac{9}{10}n}\frac{1}{n-i+1} = n\cdot\sum_{i=\frac{1}{10}n+1}^{n}\frac{1}{i} = n\left(\sum_{i=1}^{n}\frac{1}{i} - \sum_{i=1}^{\frac{1}{10}n}\frac{1}{i}\right) = n\left(H_n - H_{\frac{1}{10}n}\right) = n\left(\ln n - \ln\left(\frac{1}{10}n\right) + O(1)\right)$$

$$= n\left(\cancel{\ln n} - \left(\cancel{\ln(n)} - \ln(10)\right) + O(n)\right)$$

$$= n\ln(10) + O(n)$$

$$= \Theta(n)$$

$$\Rightarrow \underline{\underline{c = 0}}$$

$$\frac{1}{n-\frac{9}{10}n+1} = \frac{1}{\frac{1}{10}n+1}$$

$$\frac{1}{n-\left(\frac{9}{10}n-1\right)+1} = \frac{1}{\frac{1}{10}n+2}$$

$$\vdots$$

$$\frac{1}{n-1+1} = \frac{1}{n}$$

Exercise: (Doublesums)

We have blue and red beads in a pot and want to build a necklace with n beads. To do this, we randomly draw a bead from our pot and add it to the necklace. We assume that we have an infinite supply of beads and that each draw results in a red bead with probability 0.25 and a blue bead with probability 0.75. After adding n beads in sequence to our necklace (where the beads are numbered in order), we close it into a loop. We now ask how many color transitions occur in the necklace, meaning how many positions exist where the i-th bead has a different color than the $i+1$-th bead (where the $n+1$-th bead is the same as the 1st bead, since it forms a loop). Let

$$X := \text{Number of color transitions in the necklace}$$

To start, we first define indicator variables for all subproblems (which is almost always a good approach) to simplify the problem. One can verify (or look it up in a reference) that X does not follow a binomial distribution.

$$X_i := \text{Indicatorvariable for the event "There is a color transition at the i-th bead"}$$

(a) What is $\mathbb{E}[X]$?

(b) What is $\mathbb{E}[X_i \cdot X_j]$ for arbitrary $1 \leq i \leq j \leq n$? (Hint: Case Distinction)

(c) Using (b), what is $\mathrm{Var}[X]$?

a) $E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{4}\cdot\frac{3}{4} + \frac{3}{4}\cdot\frac{1}{4} = \sum_{i=1}^{n} \frac{6}{16} = \sum_{i=1}^{n} \frac{3}{8} = \frac{3}{8}n$

b) $E[X_i X_j]$:

Case: $\underline{i=j}$ $\quad E[X_i X_j] = E[X_i^2] \overset{\text{last class}}{=} E[X_i] = \frac{3}{8}$ $\quad \leftarrow n$ times

$\underline{X_i \,\&\, X_j \text{ indep.:}}$ $\quad E[X_i X_j] = E[X_i]\,E[X_j] = \frac{9}{64}$ $\quad \leftarrow \underbrace{n^2 - n - 2n}_{n^2-3n}$ times

$\underline{X_i \,\&\, X_j \text{ dep.:}}$ (beside one another)

$X_i X_{i+1}$ &

$X_i X_{i-1}$

$E[X_i X_j] = 1 \cdot \mathbb{P}[X_i X_j = 1] + 0 \cdot \mathbb{P}[X_i X_j = 0]$ $\quad \leftarrow 2n$ times, as for each i, there are two neighbours

$= \mathbb{P}[X_i X_j = 1]$

$= \mathbb{P}[X_i = 1, X_j = 1]$

$= \frac{3}{4}\cdot\frac{1}{4}\cdot\frac{3}{4} + \frac{1}{4}\cdot\frac{3}{4}\cdot\frac{1}{4}$

$= \frac{9}{64} + \frac{3}{64} = \frac{12}{64}$

c) $\mathrm{Var}[X] = E[X^2] - E[X]^2$

$= E[\sum_{i=1}^{n} X_i \sum_{j=1}^{n} X_j] - \left(\frac{3}{8}n\right)^2$

$= E[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j] - \left(\frac{3}{8}n\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} E[X_i X_j] - \frac{9}{64}n^2$

$= \left(n\cdot\frac{3}{8} + 2n\cdot\frac{12}{64} + (n^2-3n)\cdot\frac{9}{64}\right) - \frac{9}{64}n^2$