

Expectation $E[X] = \sum_{x \in \Omega(X)} x \cdot P(X=x) = \sum x \cdot P(x)$

R.v. X requirement $\text{values}(X) = \text{numbers}$

intuition
 $E[X] =$ weighted average of values weighted
 By probabilities of seeing that value
 $\frac{P(x)}{P(X=x)}$

uniform: $P(x) = \text{constant} = \frac{1}{|\Omega(X)|} = \frac{1}{n}$

$n = \# \text{ of values}$

$$E[X] = \sum_x x \cdot P(x) = \frac{1}{n} \sum_x x =$$

$$= \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{arithmetic average}$$

$E[\]$ rules $c = \text{constant}$

$$E[c \cdot X] = c E[X] \quad \text{exercise}$$

$$E[c + X] = c + E[X]$$

X, Y 2 R.V with values = numbers (even DEPENDENT)

Linearity of Expectation

$$(Th) \quad E[X + Y] = E[X] + E[Y] \quad X + Y = \text{R.V.}$$

proof $E[X + Y] = \sum_{w \in \Omega(X+Y)} w \cdot P(X+Y=w) =$

$$\sum_{w = x+y \text{ values}} (x+y) P(X+Y=x+y) =$$

$$= \sum_{x \in \Omega(X)} \sum_{y \in \Omega(Y)} (x+y) P(X=x, Y=y).$$

$$\begin{aligned}
&= \sum_x \sum_y x \cdot P(x, y) + \sum_x \sum_y y \cdot P(x, y) \\
&= \sum_x x \sum_y P(x) \cdot P(y|x) + \sum_y y \sum_x P(y) \cdot P(x|y) \\
&= \sum_x \left[x P(x) \cdot \left(\sum_y P(y|x=x) \right) \right] + \sum_y \left[y P(y) \cdot \left(\sum_x P(x|y=y) \right) \right] \\
&= \sum_x x \cdot P(x) + \sum_y y \cdot P(y) \\
&= E[X] + E[Y]
\end{aligned}$$

• $E[X + Y + Z] = E[X + Y] + E[Z] = E[X] + E[Y] + E[Z]$

• $X_1, X_2, \dots, X_n$ indicator R.V. $\begin{matrix} \nearrow 1 \\ \searrow 0 \end{matrix}$ "present"/T, a not
 $X = X_1 + X_2 + \dots + X_n = \text{count of items present}$

$E(x)$ = expected # items present

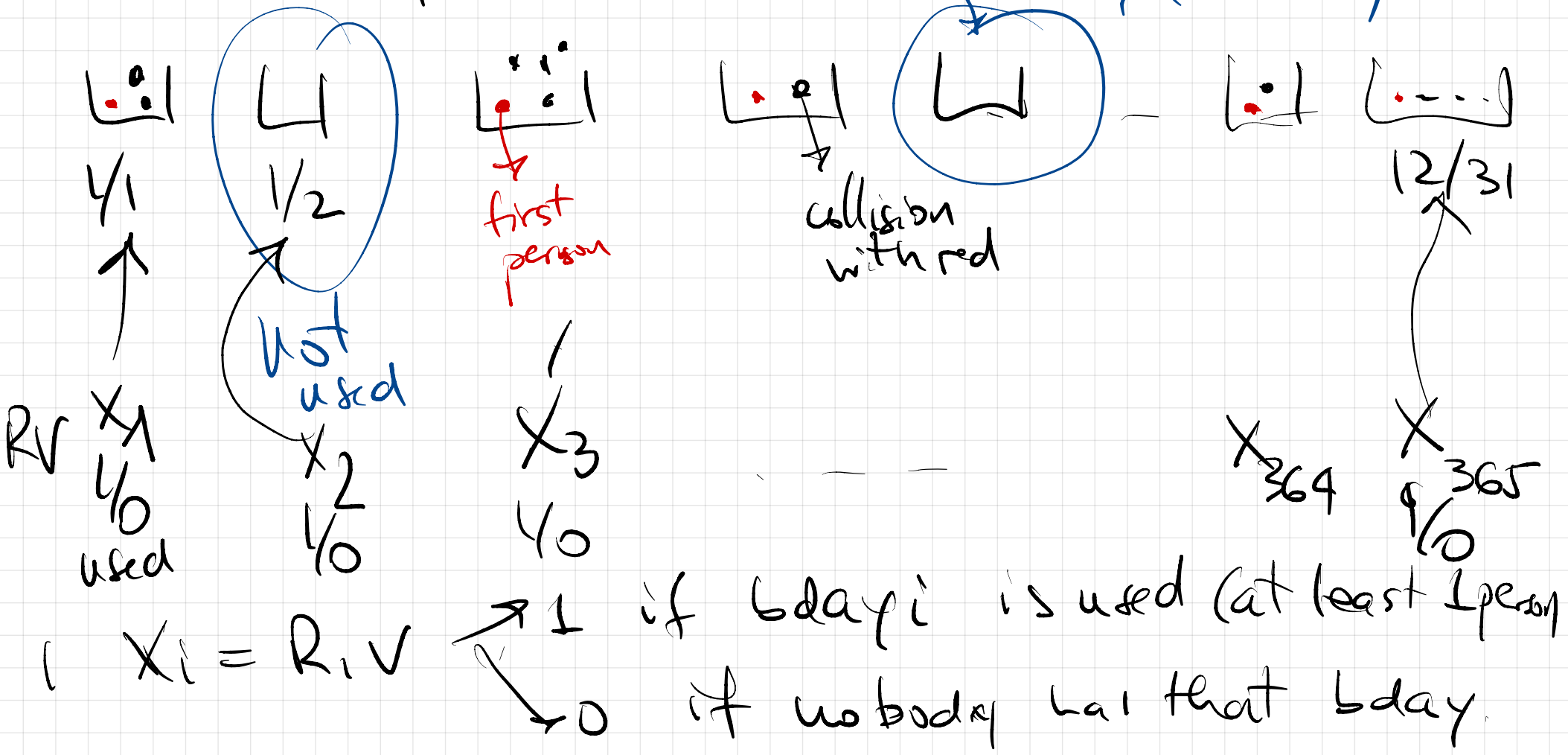
$$= E[x_1 + x_2 + \dots + x_n] = \sum_i E[x_i]$$

"expected number of ..."

Expected # of collisions? # $\leq$ days in a day with already a first-day.

ex. 1 day with 3 people $\rightarrow$ 2 collis
no people $\rightarrow$ 0 collis

X = days used (≥ 1 person that day)



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

$$E[X] = E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] + \dots + E[X_n]$$

$$E[X_i] = E[X_j] \text{ (all have the same expectation)}$$

$$E[X_1] = \text{prob}(X_1=1) \cdot 1 + \text{prob}(X_1=0) \cdot 0 = \text{prob}(X_1=1)$$

$$P(X_1=1) = 1 - \text{prob}(X_1=0) =$$

→ no body has
X₁ bday

$$1 - \left(1 - \frac{1}{365}\right)^n \rightarrow n \text{ people}$$

→ probab of 1 person not
have that bday X₁

X, Y R.V

$$E[X \cdot Y] = \sum_{w=x \cdot y} x \cdot y \text{ Prob}(X \cdot Y = w)$$

$$= E[X] \cdot E[Y] \text{ if } X, Y \text{ independent}$$

$$\neq E[X] \cdot E[Y] \text{ in general } X, Y \text{ might be dependent.}$$

coin $P_r(\text{head}) = p$ $P_r(\text{Tail}) = 1-p$ Bernoulli $\rightarrow$ Binomial distribution

$n = \# \text{ flips}$ INDEP $X = \# \text{ of heads in } n \text{ flips}$

$k = \text{target}$

$$\textcircled{1} P[X=k] \quad \frac{H/T}{F_1} \quad \frac{H/T}{F_2} \quad \frac{H/T}{F} \quad \dots \quad \frac{H/T}{F_n}$$

exact k heads $\Leftrightarrow k$ flips are "Heads" $n-k$ flips are "tails"

choose k flips $\rightarrow$ "heads" $\binom{n}{k}$ sequences H/T with k heads

prob for a fix sequence

example

k heads
prob

H T T H . . . H T . . . T

$$p (1-p) (1-p) p$$

$$p (1-p)$$

$$(1-p)$$

flips are indep

$$p^k (1-p)^{n-k}$$

$$P[X=k] = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

prob of k exact heads and exact $n-k$ tails

Sanity check: $1 = \sum_{k=0}^n \underbrace{P[X=k]}_{\text{binomial dist}} = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n \underset{1}{=}$

$x=p; y=1-p$

② $E[X] = ?$

def: $\sum_{\substack{0 \leq k \leq n \\ \text{heads}}} k \cdot P[X=k] = \sum_k k \binom{n}{k} p^k (1-p)^{n-k} = \sum_k k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$

$= np \sum_k \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$

$= np \sum_{0 \leq k-1 \leq n-1} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$

$[p + (1-p)]^{n-1}$
exercise

$= np$

$$\textcircled{2} E[X] = E[X_1 + X_2 + \dots + X_n] \quad \text{1 binary Rv per flip}$$

$$X_i = \begin{cases} 1 & \text{if flip } i = H \\ 0 & \text{if flip } i = T \end{cases}$$

$$= n \cdot E[X_1] = n \cdot \text{prob}(X_1 = 1) = n \cdot p$$

$P(H) = p$
Heads = Success

Flip coin every day
Tail = Failure $P(Tails) = 1-p$

R.V. $X = \#$ flips until first success.

distribution (x) ?

$E(X) = ?$

$$P(X=1) = p$$

$$P(X=2) = (1-p) \cdot p$$

$$P(X=3) = (1-p)(1-p) \cdot p = (1-p)^2 \cdot p$$

$$P(X=k) = \underbrace{(1-p) \dots (1-p)}_{k-1 \text{ times}} \cdot \underbrace{p}_{\text{success}} = (1-p)^{k-1} \cdot p$$

$$P(X=10^8)$$

Sanity check: $\sum_{k=1}^{\infty} P\{X=k\} = \sum_{k=1}^{\infty} p(1-p)^{k-1}$

$\sum \text{probab} = p \sum_{k=1}^{\infty} (1-p)^{k-1}$ $x = 1-p$

$$= p \cdot (x^0 + x^1 + x^2 + \dots + x^{108} + \dots)$$

$$= p \cdot \frac{1}{1-x}$$

$$= p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1$$

$E[x]$ = expected # of trials until first success
(average)

$$\checkmark \Rightarrow = \sum_{k=0}^{\infty} k \cdot P[x=k] = \sum_{k=1}^{\infty} k \cdot \boxed{P(1-p)^{k-1}}$$

Geometric

TELESCOPE $\checkmark$ summation

$$\checkmark(1-p) =$$

$$\sum_{k=1}^{\infty} k \cdot p(1-p)^k$$

$$\checkmark - \checkmark(1-p) = \underbrace{p}_{k=1} + \sum_{k=2}^{\infty} \underbrace{k p}_{k=2} (1-p)^{k-1} - \sum_{k=2}^{\infty} \underbrace{(k-1)p}_{\text{indexed by } k-1} (1-p)^{k-1}$$

$$= p + \sum_{k=2}^{\infty} \underbrace{[k p - (k-1)p]}_p (1-p)^{k-1}$$

$$= p + \sum_{k=2}^{\infty} p (1-p)^{k-1}$$

$$= \sum_{k=1}^{\infty} p(1-p)^{k-1} = 1 \quad \text{from before}$$

$$\text{So } v - v(1-p) = 1$$

$$vp = 1$$

$$E[X] = v = \frac{1}{p}$$

$$\text{Example } p = \text{prob}(\text{success}) = \frac{1}{5}$$

$$\Rightarrow E[\# \text{ trials to success}] = \frac{1}{1/5} = 5$$

Balls into Bins m balls are indep thrown into one of the k bins, uniformly.

- ① What is the expected # balls falling into a bin?
- ② What is the expected # of empty bins?

① Define R.V. $X_{ij} = \begin{cases} 1 & \text{if ball } i \rightarrow \text{bin } j \\ 0 & \text{if not.} \end{cases}$ Total $m \times k$ R.V.

$$E[\# \text{ of balls in bin } t] = E\left[\sum_{i=1}^m X_{it}\right] = \sum_{i=1}^m E[X_{it}] = m \cdot \frac{1}{k}$$

pick at $1 \leq t \leq k$

"1" for every ball $\rightarrow$ bin t

$$E[X_{it}] = 1 \cdot p(X_{it}=1) + 0 \cdot p(X_{it}=0) = p(X_{it}=1) = \frac{1}{k}$$

prob of ball $i \rightarrow$ bin t
unif over k bins

② Fix a bin t , $1 \leq t \leq k$.
R.V. $Y_t = \begin{cases} 1 & \text{if all balls miss bin } t \Rightarrow \text{empty} \\ 0 & \text{if not} \Rightarrow \text{not empty} \end{cases}$

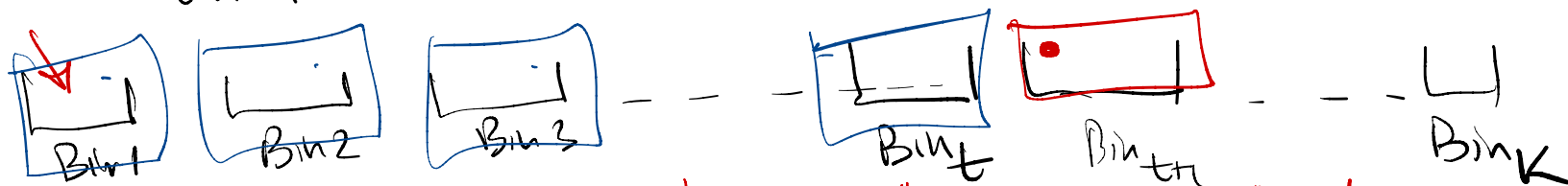
$$P[Y_t=1] \xrightarrow[m \text{ throws}]{\text{indep}} \prod_{i=1}^m \boxed{\text{ball } i \text{ miss bin } t} = \prod_{i=1}^m P[X_{it}=0] = \prod_{i=1}^m \left(1 - \frac{1}{k}\right)$$

$$= \left(1 - \frac{1}{k}\right)^m$$

$$E[\# \text{ empty bins}] = E\left[\sum_{t=1}^k Y_t\right] = \sum_{t=1}^k E[Y_t] = k \cdot \left(1 - \frac{1}{k}\right)^m \approx k \cdot e^{-m/k}$$

optional: "Coupon Collector PB"

3* Say k is fixed (#bins), but m is a random variable: we throw m balls until all the bins are non-empty. What is the $E[m]$?



Bins are now in "hit em + order" $\rightarrow$ coupon collected order.

- First ball $\rightarrow$ Bin 1.

#balls going to bin 1 X_1

- Next bin that gets its first ball Bin 2 #balls until we hit bin 3 X_2

- Next empty bin gets hit

Bin 3 #balls until we hit bin 4 X_3

- Next Bin 4

#balls until we hit bin 5 X_4

- Bin t #balls until we hit (first time) bin t X_t

Bin t+1

#balls until we hit bin t+1 X_{t+1}

$X_t =$ #balls after hitting bin t (first time) until we hit bin t+1 (first time)

$$p(\text{*ball} \text{ to hit bin } t+1) = 1 - p(\text{*ball falls into one of the non-empty bins})$$

$$= 1 - \frac{t}{k} \quad \text{"chance of success" fixed}$$

⇒ geometric distribution
(exercise)

$x \in [0, 1]$

$$1 + x + x^2 + x^3 + \dots = \frac{1}{1-x}$$

$$= \frac{1}{\text{success rate}} = \frac{1}{1 - \frac{1}{k}} = \frac{k}{k-t+1}$$

$$E[X_t] = E[X = \text{geometric dist}]$$

Success = $1 - \frac{1}{k}$

$$\# \text{ balls} = 1 + \sum_{t=1}^k x_t = \# \text{ balls till bin 1} \Rightarrow 1$$

balls till bin 2

$$= \sum_{t=1}^k \frac{k}{k-t+1}$$

$$= k \sum_{t=1}^k \frac{1}{k-t+1}$$

balls till bin k

$$= k \cdot \sum_{t=1}^k \frac{1}{t} \approx k \cdot H_k$$

Harmonic H_k

$$H_k = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{k}$$

$$\text{Variance (R.V } x) = \text{var}(x) = \text{avg}(\text{distance-to-mean})^2$$

$$= \text{avg} \left(x - \underset{\text{mean}}{E[x]} \right)^2 = E[(x - E[x])^2]$$

$E[x]$ = not random
value = μ

x = R.V. random

$$\text{var}(x) = E[(x - \mu)^2] = E[x^2 - 2x\mu + \mu^2] =$$

$$= E[x^2] - 2\underset{\mu}{\mu}E[x] + \underset{E[\mu^2]}{\mu^2}$$

$$= E[x^2] - 2\mu^2 + \mu^2 = E[x^2] - \underset{E[x]}{\mu^2}$$

$$= E[x^2] - (E[x])^2$$

$$\text{var}(\underset{x}{\text{die roll}})$$

$$\mu = E[x] = 3.5 = \frac{7}{2}$$

$$= \text{avg}((x - \mu)^2) =$$

$$\frac{1}{6} \left((1 - \frac{7}{2})^2 + (2 - \frac{7}{2})^2 + (3 - \frac{7}{2})^2 + (4 - \frac{7}{2})^2 + (5 - \frac{7}{2})^2 + (6 - \frac{7}{2})^2 \right)$$

$$= \frac{70}{24} ?$$

Coin $p(\text{heads}) = p = \frac{1}{2}$
 $x = 1$ if heads

$$p(\text{tails}) = 1 - p = \frac{1}{2}$$

$$\mu = E[x] = 1 \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{2}{3} =$$

$$\begin{aligned} \text{var}[x] &= \left(1 - \frac{2}{3}\right)^2 \cdot p(x=1) + \left(0 - \frac{2}{3}\right)^2 \cdot p(x=0) = \\ &= \left(\frac{1}{3}\right)^2 \cdot \frac{2}{3} + \left(\frac{2}{3}\right)^2 \cdot \frac{1}{3} = \frac{6}{27} \end{aligned}$$

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{\omega \in \Omega} X(\omega)Y(\omega)\Pr(\omega) \\
&= \sum_x \sum_y xy \cdot \Pr(X = x \text{ and } Y = y) \\
&= \sum_x \sum_y xy \cdot \Pr(X = x)\Pr(Y = y) \\
&= \left(\sum_x x \cdot \Pr(X = x) \right) \left(\sum_y y \cdot \Pr(Y = y) \right) \\
&= \mathbb{E}(X)\mathbb{E}(Y),
\end{aligned}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2.$$

Consequently, $\text{Var}(X) \leq E(X^2)$.

Proof. Using the linearity of expectation and the fact that the expectation of a constant is itself, we have

$$\begin{aligned}\text{Var}(X) &= E(X - E(X))^2 \\ &= E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2\end{aligned}$$

Proposition 6.11. *Given a discrete probability space (Ω, Pr) , for any random variable X and Y , if X and Y are independent, then*

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Proof. Recall from Proposition 6.9 that if X and Y are independent, then $E(XY) = E(X)E(Y)$. Then, we have

$$\begin{aligned}E((X + Y)^2) &= E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) \\ &= E(X^2) + 2E(X)E(Y) + E(Y^2).\end{aligned}$$

Using this, we get

$$\begin{aligned}\text{Var}(X + Y) &= E((X + Y)^2) - (E(X + Y))^2 \\ &= E(X^2) + 2E(X)E(Y) + E(Y^2) - ((E(X))^2 + 2E(X)E(Y) + (E(Y))^2) \\ &= E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 \\ &= \text{Var}(X) + \text{Var}(Y),\end{aligned}$$