Definition of a "fair game" is that $\mu = 0$.

If the P(traffic) = 0.3 and it takes 7min without traffic and 12 min with traffic. What is your expected trip time?

$$W \sim \begin{cases} 7_{min} & \text{wp } 0.7 \\ 12_{min} & \text{wp } 0.3 \end{cases}, \quad E[W] = 7_{min} \cdot 0.7 + 12_{min} \cdot 0.3 = \underline{8.5_{min}}$$

The taxi company charges you \$0.40/min. Let's model the charges due to time, B.

$$B = \$0.40/\text{min}\ W \sim \begin{cases} \$2.80 & \text{wp } 0.7 \\ \$4.80 & \text{wp } 0.3 \end{cases}, \quad E[B] = \$2.80 \cdot 0.7 + \$4.80 \cdot 0.3 = \$3.12$$

Note: $\underbrace{\$0.40/\text{min} \cdot 8.5\text{min}}_{E[B] = \$0.40/\text{min}\, E[W]}$ $''$

This is a "transformation" or "function" of a rv which produces another rv.

$$E[Y] \overset{\downarrow}{=} E[aX] = \sum_{x \in Supp[X]} a x\, p(x) = a \overbrace{\sum_{x \in Supp[X]} x\, p(x)}^{E[X]} = a\, E[X].$$

There is always a base fee of \$3. What is the bill for the time with this base fee?

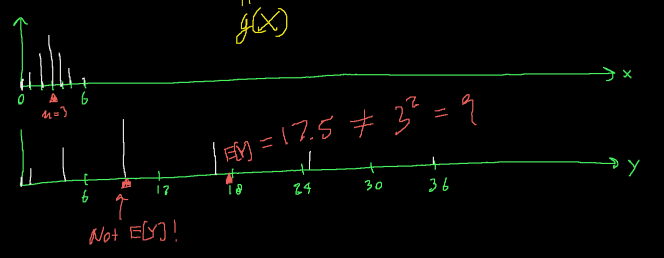
$$T = B + \$3 \sim \begin{cases} \$5.80 & \text{wp } 0.7 \\ \$7.80 & \text{wp } 0.3 \end{cases}, \quad E[T] = \$5.80 \cdot 0.7 + \$7.80 \cdot 0.3 = \$6.12$$

Note: $E[T] = E[B] + \$3$ $''$

let Y = X + c where $c \in \mathbb{R}$ constant and X is a rv, then E[Y] = E[X] + c.

$$E[Y] \overset{\downarrow}{=} E[X + c] = \sum (x+c)\, p(x) = \underbrace{\sum x\, p(x)}_{E[X]} + \sum c\, p(x) = E[X] + c \underbrace{\sum p(x)}_{1} = E[X] + c.$$

$$X \sim Bin(6, \tfrac{1}{2}), \quad Y = \underset{g(X)}{\underset{\parallel}{X^2}} \qquad E[Y] \neq E[X]^2 \quad \text{No. Not generally.}$$



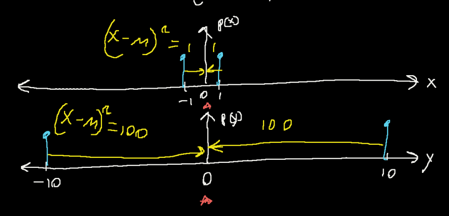
E[Y] is not calculable based on a formula. You have to do it manually i.e. $\mu_Y = \sum_{y=0}^{n} y^2 \binom{n}{y} p^y (1-p)^{n-y}$

$$\Rightarrow E[g(X)] = \sum_{x \in Supp[X]} g(x)\, p(x)$$

proof is beyond scope of course because it requires you to dig in using $\omega \in \Omega$ which we decided to ignore a few lectures ago.

$$\Rightarrow E[g(X)] \neq g(E[X]) \quad \text{Not generally}$$

$$X \sim \text{Rademacher} := \begin{cases} +1 & \text{wp } 0.5 \\ -1 & \text{wp } 0.5 \end{cases}, \quad Y = 10X \sim \begin{cases} +10 & \text{wp } 0.5 \\ -10 & \text{wp } 0.5 \end{cases}$$



$E[X] = 0$

$E[Y] = E[10X]$
$= 10\,E[X]$
$= 0$

$E[X] = E[Y] \not\Rightarrow p(x) = p(y)$

$p(x) = p(y) \Rightarrow E[X] = E[Y]$

It is clear that Y is more "spread out" from its expectation than X. The values of y are thus more "variable" than the values of x or have higher "variance" or "deviance" than x.

Is there a metric you can calculate that will allow for these comparisons?

Let's discuss "error functions" (or "loss functions"). These must be ≥ 0 for all values of $x \in Supp[X]$. These measure "distance from pivot".

$e(x, \mu) = |x - \mu|$ This is called L1 loss and it is well-used.
$e(x, \mu) = (x - \mu)^2$ This is called L2 loss and it is the default. We will use L2 loss for this class.

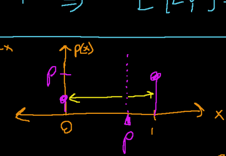
We then define the variance of a rv as the expected L2 loss.
Let $L := (X - \mu)^2$.

$$\sigma^2 := Var[X] := E[L] = E[(X - \mu)^2] = \sum_{x \in Supp[X]} (x - \mu)^2\, p(x)$$

$$L_X \sim \{ 1^2 \text{ wp } 1 \Rightarrow E[L_X] = 1 = Var[X]$$

$$L_Y \sim \{ 10^2 \text{ wp } 1 \Rightarrow E[L_Y] = 100 = Var[Y]$$

$X \sim Bern(p) = p^x (1-p)^{1-x}$,

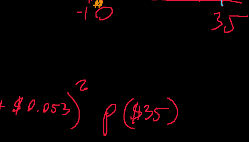


$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$

$$\begin{aligned} Var[X] = E[(X - p)^2] &= \sum_{x \in \{0,1\}} (x-p)^2 p^x (1-p)^{1-x} \\ &= (0-p)^2 \cancel{p^0} (1-p)^{1-0} + (1-p)^2 p^1 \cancel{(1-p)^{1-1}} \\ &= (-p)^2 (1-p) + (1-p)^2 p \\ &= p^2(1-p) + (1 - 2p + p^2)\, p \\ &= \cancel{p^2} - \cancel{p^3} + p - \cancel{2}p^2 + \cancel{p^3} \\ &= p - p^2 = p(1-p) \end{aligned}$$

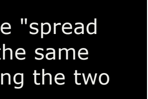
Recall roulette in America. Bet on lucky #7. Payout is 35:1

$$X \sim \begin{cases} \$35 & \text{wp } \frac{1}{38} \\ -\$1 & \text{wp } \frac{37}{38} \end{cases} \qquad E[X] = \ldots = -\$0.053$$



$$\begin{aligned} Var[X] &= \sum_{x \in \{-\$1, \$35\}} (x - -\$0.053)^2\, p(x) \\ &= (-\$1 + \$0.053)^2\, p(-\$1) + (\$35 + \$0.053)^2\, p(\$35) \\ &= 0.897\,\$^2\, \frac{37}{38} + 1228.713\,\$^2\, \frac{1}{38} \\ &= 33.207\,\$^2 \end{aligned}$$

Bet on black. Payout is 1:1



$$X \sim \begin{cases} \$1 & \text{wp } \frac{18}{38} \\ -\$1 & \text{wp } \frac{20}{38} \end{cases} \qquad E[X] = \ldots = -\$0.053$$

$$Var[X] = (\$1 - -\$0.053)^2 \frac{18}{38} + (-\$1 - -\$0.053)^2 \frac{20}{38} = 0.997\,\$^2$$

Variance seems to do its job: we can now measure the "spread out"ness of a rv. And we can compare rv's that have the same mean on their degree of "spread out"ness by computing the two variances and comparing.

But... the unit of variance does not make sense! We would like its unit to be the same as the unit that x is measured on. To convert variance to the same unit as x i.e. we take the "deviation" and make it "standard" we just take the square root and define the "standard deviation" of a rv:

$\sigma := SD[X] := \sqrt{Var[X]}$

$\bar{X}_{\#7}$ vs $\bar{X}_{Black}$

by law of large numbers, both these averages converge to the same mean μ = -\$0.053. But which when "converges faster"? The rv with the lower variance, i.e. betting on Black.