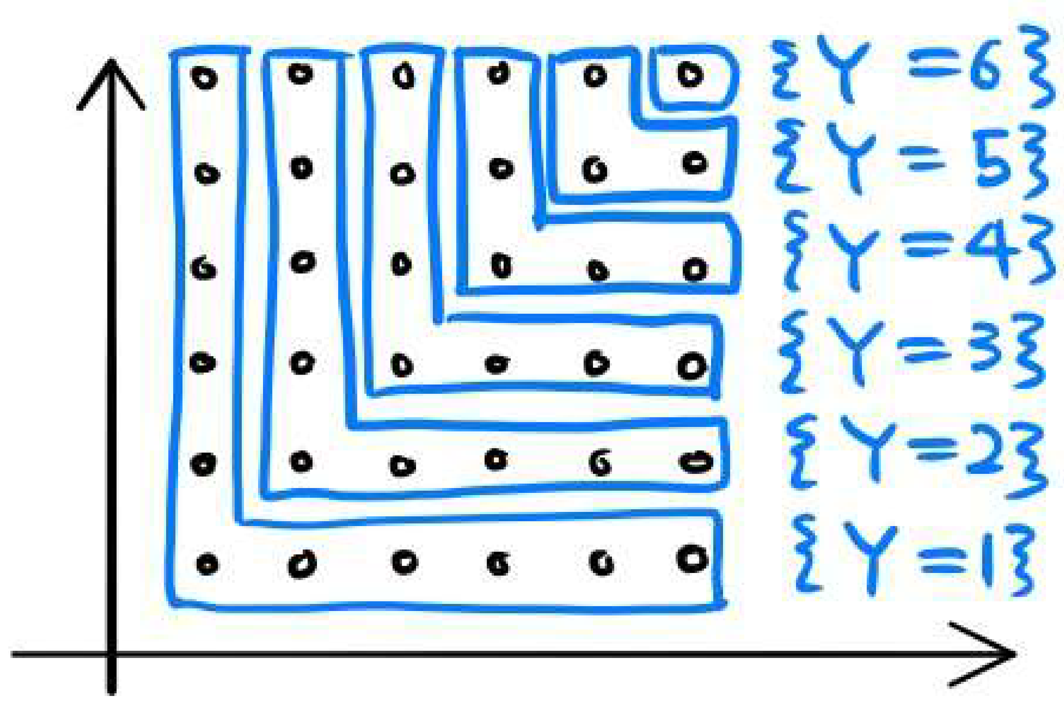
Review)

- If X is a RV of discrete type with PMF p_X, then
$$E(X) = \sum_{x \in S_X} x\, p_X(x).$$
- LOTUS tells that
$$E(u(X)) = \sum_{x \in S_X} u(x) p_X(x).$$
- $E(\cdot)$ satisfies
$$E(c) = c \quad \text{and} \quad E(c_1 X_1 + \cdots + c_n X_n) = c_1 E(X_1) + \cdots + c_n E(X_n)$$
for constants $c, c_1, \cdots, c_n$ and RVs $X_1, \cdots, X_n$.

$\boxed{\text{Ex}}$ For each constant b,
$$\begin{aligned} E[(X-b)^2] &= E[X^2 - 2bX + b^2] \\ &= E[X^2] - 2b E[X] + b^2 \\ &= (E[X]-b)^2 + (E[X^2] - E[X]^2). \end{aligned}$$
This is minimized when $b = E[X]$ with the minimum value $E[X^2] - E[X]^2$.

$\boxed{\text{Ex}}$ Roll two 6-faced dices and let
$$X_1 = [\text{outcome of } 1^{st} \text{ dice}]$$
$$X_2 = [\text{——— " — } 2^{nd} \text{ dice}]$$
also, set
$$Y = \min\{X_1, X_2\}, \quad Z = \max\{X_1, X_2\}, \quad W = X_1 + X_2.$$
Then

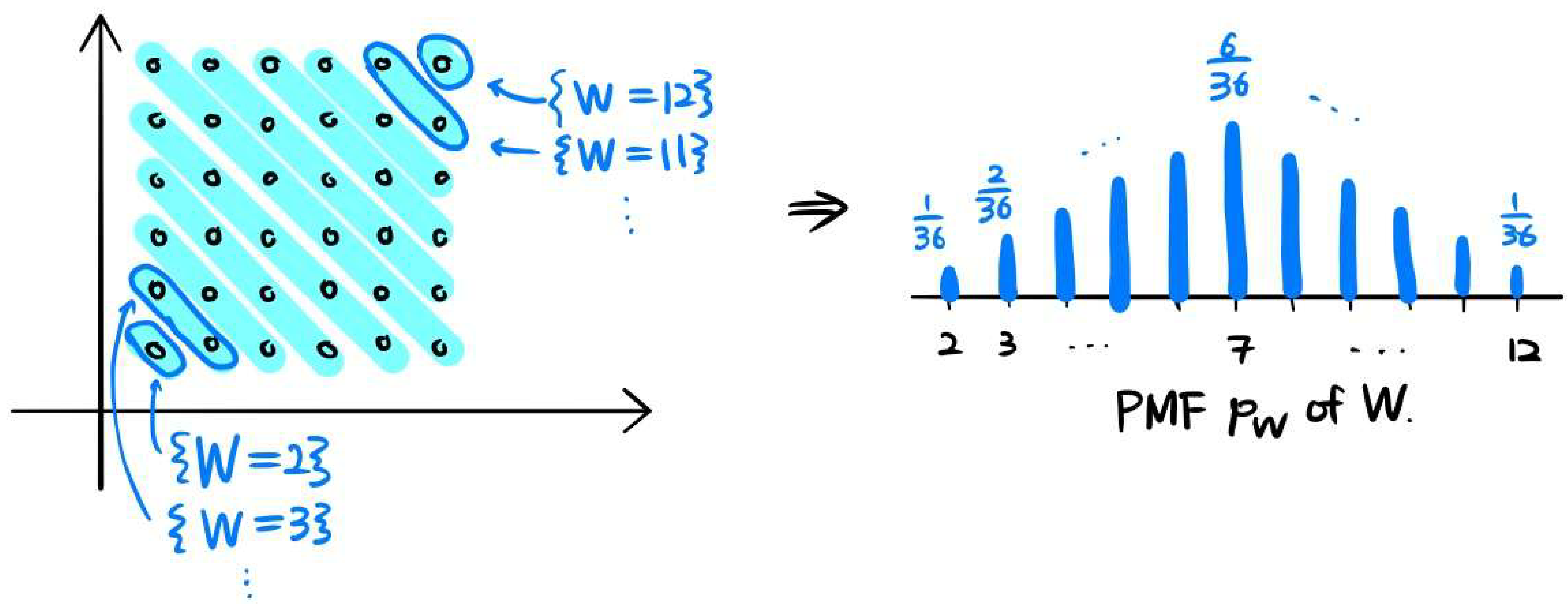


$$\Rightarrow \quad p_Y(k) = \frac{13-2k}{36} \quad \text{for } k = 1, 2, \cdots, 6.$$
$$\Rightarrow \quad E(Y) = \sum_{k=1}^{6} k \cdot p_Y(k) = \sum_{k=1}^{6} \frac{k(13-2k)}{36} = \frac{91}{36}.$$

Similar argument shows $E(Z) = \frac{161}{36}$. Alternatively, using the identity
$$X_1 + X_2 = Y + Z,$$
We get
$$E(Z) = E(X_1 + X_2 - Y) = E(X_1) + E(X_2) - E(Y).$$

Since $E(X_1) = \sum_{x=1}^{6} x \cdot \frac{1}{6} = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}$ and $E(X_2) = \frac{7}{2}$ likewise, we get $E(Z) = \frac{7}{2} + \frac{7}{2} - \frac{91}{36} = \frac{161}{36}$ as we have already confirmed.

Also, $E(W) = E(X_1) + E(X_2) = 7$. This can also be directly checked using



$\square$

Ex (Geometric distribution) Assume:

- A random expr. yields $\begin{cases} \text{SUCCESS with prob } p \in (0,1), \\ \text{FAILURE with prob } q = 1-p. \end{cases}$
- Repeat this expr. indefinitely.
- $X = [\#$ of trials until the 1^{st} SUCCESS$]$.

Then

- $S_X = \{1, 2, 3, \cdots\}$.
- For each $x \in S_X$, i.e., for $x = 1, 2, 3, \cdots$,

$$\begin{aligned} p_X(x) &= P(X = x) \\ &= P((x-1) \text{ FAILURE then 1 SUCCESS}) \\ &= \underbrace{q \cdots q}_{(x-1) \text{ times}} \cdot p = q^{x-1} p. \end{aligned}$$

(A sanity check: $\sum_{x \in S_X} p_X(x) = \sum_{x=1}^{\infty} q^{x-1} p = \frac{p}{1-q} = 1$. ☺)

- $E(X) = \sum_{x=1}^{\infty} x \cdot p_X(x) = \sum_{x=1}^{\infty} x q^{x-1} \cdot p$.

Using the formula

$$\frac{1}{(1-z)^2} = \frac{d}{dz} \frac{1}{1-z} = \frac{d}{dz}(1 + z + z^2 + \cdots) = (1 + 2z + 3z^2 + \cdots) = \sum_{n=1}^{\infty} n z^{n-1},$$

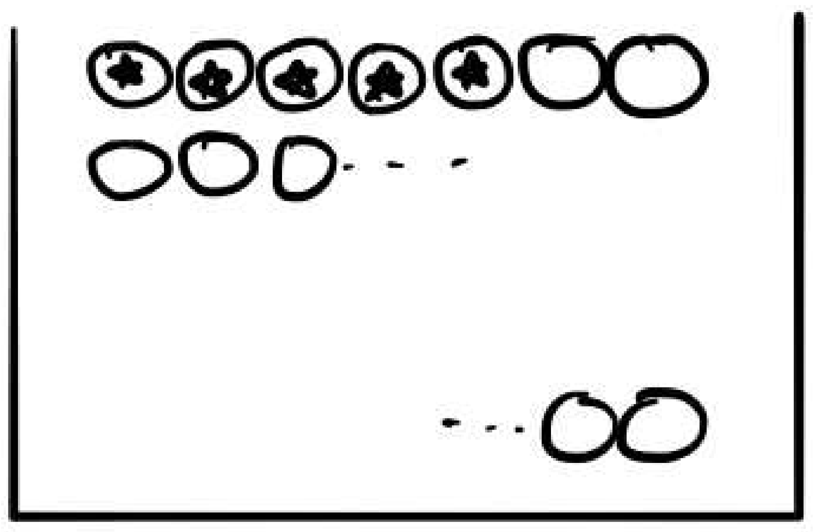
We get $E(X) = \frac{p}{(1-q)^2} = \frac{1}{p}$.

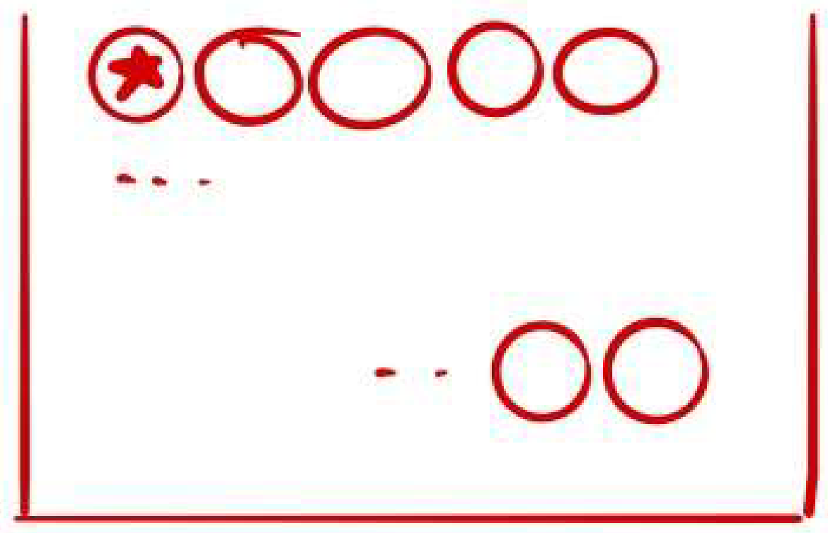
$\square$

Ex

- Box 1 : 69 white balls, 5 of them are labeled with ★.

  Box 2 : 26 red balls, one of them is labeled with ★.





- Pick 5 white balls from Box 1 w/o replacement, Pick 1 red ball from Box 2

- The player wins a prize according to the following table :

Match	Prize	# of ways (out of total $\binom{69}{5}\binom{26}{1}$)
⊛⊛⊛⊛⊛ + ⊛	343m	$\binom{5}{5}\binom{64}{0}\binom{1}{1}\binom{25}{0} = 1$.
⊛⊛⊛⊛⊛	1m	$\binom{5}{5}\binom{64}{0}\binom{25}{1} = 25$
⊛⊛⊛⊛ + ⊛	50k	$\binom{5}{4}\binom{64}{1}\binom{1}{1}\binom{25}{0} = 320$
⊛⊛⊛⊛	100	$\binom{5}{4}\binom{64}{1}\binom{25}{1} = 1{,}600$
⊛⊛⊛ + ⊛	100	⋮ 8,000
⊛⊛⊛	7	504,000
⊛⊛ + ⊛	7	416,640
⊛ + ⊛	4	3,176,880
⊛	4	7,624,512.

- The expected prize is then

$$343{,}000{,}000 \times \frac{1}{292{,}210{,}338} + 1{,}000{,}000 \times \frac{25}{292{,}210{,}338} + \cdots$$

$$= 1.49372\cdots$$

It is unfair to pay \$2 to enter this game. ∎

Section 2.3. Special Mathematical Expectations

GOAL Study the quantities derived from RV which have certain statistical meanings.

- $E(X)$ of X is often called the mean of X (or of its distribution) and denoted as μ.
  ← mu

- The "second moment of X about the mean μ"

$$E[(X-\mu)^2]$$

  is called the variance of X (or of its distribution) and denoted as $\mathrm{Var}(X)$.

- The square root

$$\sqrt{E[(X-\mu)^2]}$$

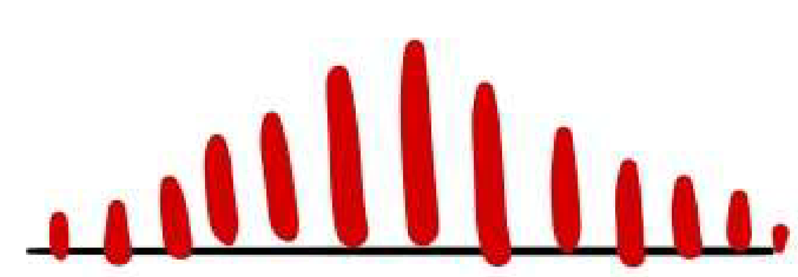
  is called the standard deviation of X and denoted as σ. Accordingly,

sigma

$$\mathrm{Var}(X) = \sigma^2.$$

  σ (as well as $\mathrm{Var}(X)$) is a measure of how the distribution of X is dispersed:

"

Var [] < Var []

"

$\boxed{\text{PROP}}$ $\quad \mathrm{Var}(X) = E[X^2] - \mu^2 \quad$ (where $\mu = E[X]$).

Pf) $$\mathrm{Var}(X) \overset{\text{by def.}}{=} E[(X-\mu)^2]$$
$$= E[X^2 - 2\mu X + \mu^2]$$
$$= E[X^2] - 2\mu E[X] + \mu^2$$
$$= E[X^2] \underbrace{- 2\mu\cdot\mu + \mu^2}_{= -\mu^2}$$

☺

$\boxed{\text{Ex}}$ (Uniform distribution) Let X have the unif. dist. over $S = \{1, \cdots, m\}$, i.e.,

$$p_X(x) = \begin{cases} \frac{1}{m} & \text{for } x = 1, \cdots, m \\ 0 & \text{o/w.} \end{cases}$$

Then

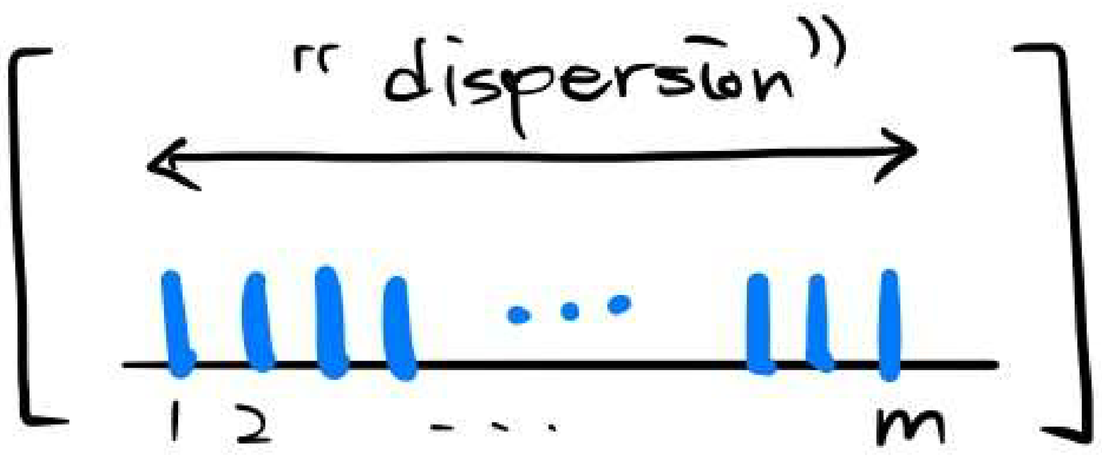
- The mean $\mu = E(X)$ is
$$\mu = \sum_{x\in S} x \cdot p_X(x) = \sum_{x=1}^{m} \frac{1}{m}\cdot x = \frac{1}{m}\cdot\frac{m(m+1)}{2}$$
$$= \frac{m+1}{2}.$$

- The 2^{nd} moment $E[X^2]$ is
$$E[X^2] = \sum_{x\in S} x^2 p_X(x) = \sum_{x=1}^{m} \frac{1}{m}\cdot x^2 = \frac{1}{m}\cdot\frac{m(m+1)(2m+1)}{6}.$$
$$= \frac{(m+1)(2m+1)}{6}$$

So

$$\sigma^2 = E[X^2] - \mu^2 = \frac{m^2-1}{12}.$$

- Indeed,

Var 

$= \dfrac{m^2-1}{12}$

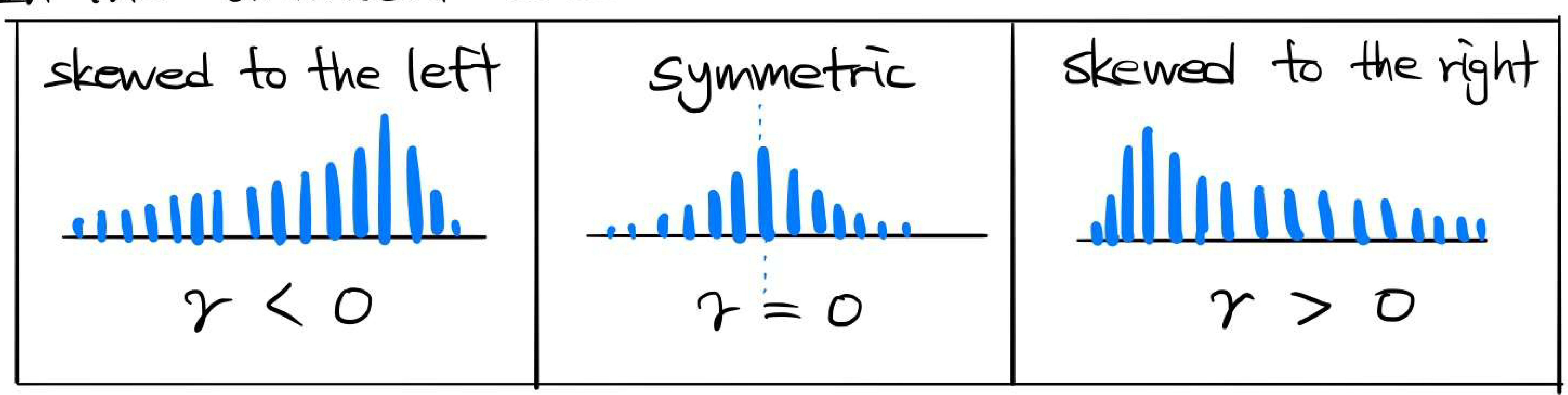
and large $Var(X)$ implies more dispersion. $\square$

$\boxed{\text{PROP}}$ (1) $E(aX+b) = aE(X)+b$, i.e., $\mu_{aX+b} = a\mu_X + b$.
(2) $Var(aX+b) = a^2 Var(X)$, i.e., $\sigma_{aX+b} = |a|\sigma_X$.

Pf) (1) is obvious from the linearity of $E(\cdot)$.

(2)
$$
\begin{aligned}
Var(aX+b) &= E[(aX+b-E(aX+b))^2] \\
&= E[(aX-aE(X))^2] \\
&= a^2 E[(X-E(X))^2] \\
&= a^2 Var(X).
\end{aligned}
$$
$\square$

- $E(X^r) = \sum_{x \in S} x^r p(x)$ is called the r-th moment of X.
- $\gamma := \dfrac{E[(X-\mu)^3]}{\sigma^3}$ is called the index of skewness of X.
- In the "unimodal" case :



$\boxed{\text{DEF}}$ Let X: RV of discrete type. Assume

$$E(e^{tX}) = \sum_{x \in S_X} e^{tx} \cdot p_X(x)$$

is finite near the origin $t=0$. Then

$$M_X(t) := E(e^{tX})$$

is called the moment-generating function (MGF) of X (or of the distribution of X).

- If we enumerate $S_X = \{b_1, b_2, \cdots\}$, then

$$M_X(t) = e^{tb_1} p_X(b_1) + e^{tb_2} p_X(b_2) + \cdots.$$

$\Rightarrow$ (1) Reading out coef. of e^{tb} in $M_X(t)$ determines the value of $p_X(b)$.

(2) So, M_X uniquely determines PMF of X.

$\boxed{\text{Ex}}$ Assume $M_X(t) = \frac{1}{6}(3e^{t} + 2e^{2t} + e^{3t})$. Then

- Support of X is $S_X = \{1, 2, 3\}$
- $p_X(1) = [\text{coef. of } e^t \text{ in } M_X(t)] = \frac{3}{6}$

  $p_X(2) = [\text{—}''\text{—} \ e^{2t} \ \text{—}''\text{—}] = \frac{2}{6}$

  $p_X(3) = \frac{1}{6}$.

$\boxed{\text{Ex}}$ Recall: $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ for any $x \in \mathbb{R}$. So, if

$$M_X(t) = e^{\lambda(e^t - 1)}, \qquad (\lambda > 0)$$

then

$$M_X(t) = e^{\lambda e^t} \cdot e^{-\lambda}$$
$$= \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} e^{-\lambda}$$
$$= \sum_{n=0}^{\infty} \left(\frac{\lambda^n}{n!} e^{-\lambda}\right) e^{tn}.$$

$$\Rightarrow P_X(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{if } x = 0, 1, 2, \cdots \\ 0 & \text{o/w.} \end{cases}$$

$\boxed{\text{PROP}}$ $M^{(r)}(0) = E(X^r)$ is the r-th moment of X.

Pf) $M^{(r)}(t) = \left(\frac{d}{dt}\right)^r M(t) = \left(\frac{d}{dt}\right)^r \sum_{x \in S} e^{tx} p(x)$

$$= \sum_{x \in S} x^r e^{tx} p(x) = E(X^r e^{tX}).$$

Plug $t = 0$. $\square$

$\boxed{\text{Ex}}$ If X has geom. dist. w/ PMF

$$P_X(x) = q^{x-1} p \quad \text{for } x = 1, 2, 3, \cdots,$$

then

$$M_X(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx} q^{x-1} p = \frac{pe^t}{1 - qe^t},$$

provided $qe^t < 1$ ($\Longleftrightarrow$ $t < -\ln q$). Then

$$E(X) = M'(0) = \frac{pe^t}{(1-qe^t)^2}\Big|_{t=0} = \frac{1}{p}$$

and

$$E(X^2) = M''(0) = \left.\frac{pe^t(1+qe^t)}{(1-qe^t)^3}\right|_{t=0} = \frac{p(1+q)}{(1-q)^3}.$$

So

$$Var(X) = E(X^2) - (E(X))^2 = \frac{q}{p^2}. \qquad \square$$