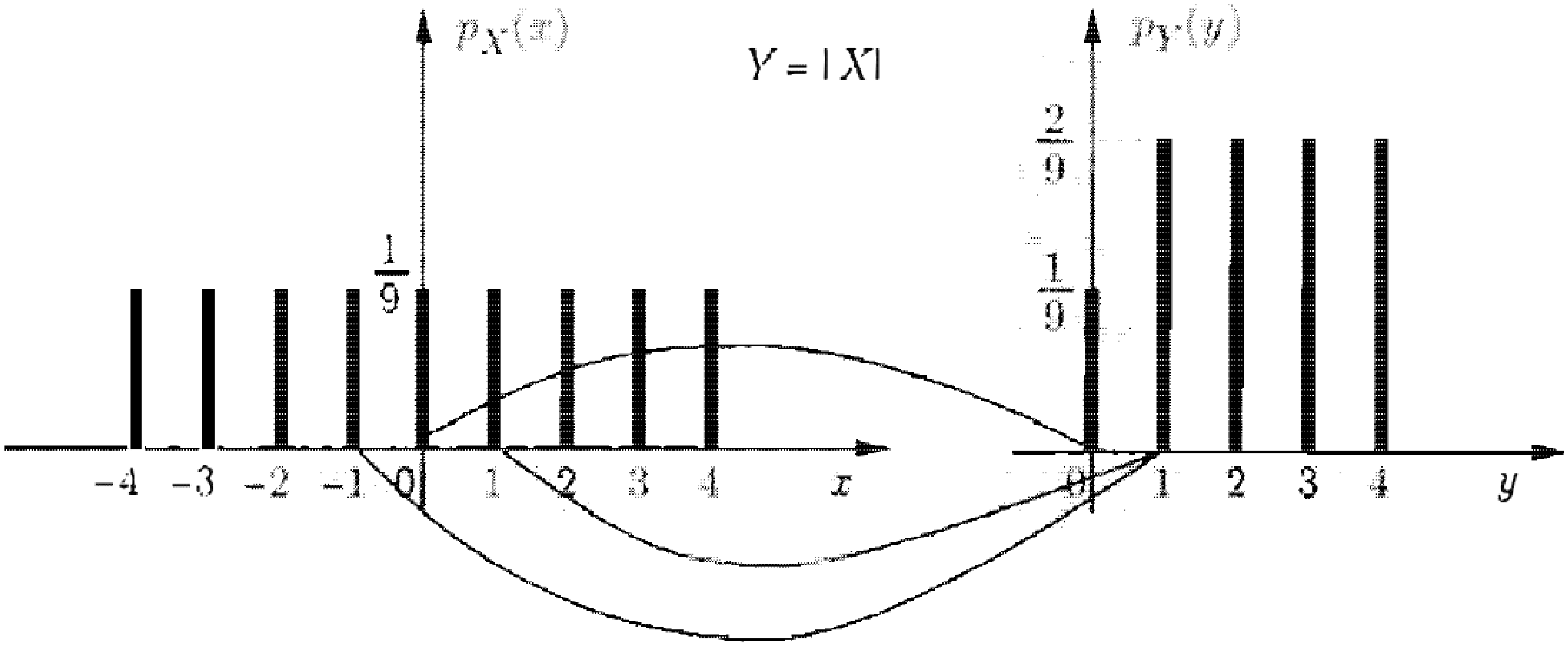
# Function of random variables

- Consider $Y = |X|$ where $X$ is the outcome of an experiment where an integer is chosen uniformly from $-4$ to 4.
- $p_X(x) = \frac{1}{9}$ for $x \in \{-4, -3, \ldots, 3, 4\}$.
- What is the range $\Omega'$ for $Y$? $\Omega' = \{0, \ldots, 4\}$.
- What is $p_Y(2)$?
- $p_Y(2) = \sum_{\{x:|x|=2\}} p_X(x) = p_X(-2) + p_x(2) = \frac{2}{9}$.

# Function of random variables

- $p_Y(2) = \sum_{\{x:|x|=2\}} p_X(x) = p_X(-2) + p_x(2) = \frac{2}{9}$.




Suppose $Y = g(X)$ and $X$ is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

# Function of random variables

Suppose $Y = g(X)$ and $X$ is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

Proof:

- Then what is $p_Y(y)$?
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : Y(\omega) = y\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : g(X(\omega)) = y\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) = g^{-1}(y)\}$.
- Is there a problem with the text in red?
- Is $g^{-1}(y)$ a value or a set?

# Function of random variables

> Suppose $Y = g(X)$ and $X$ is discrete with pmf $p_X(\cdot)$. Then $p_Y(y) = \sum_{\{x:g(x)=y\}} p_X(x)$.

Proof Continued:

- $g^{-1}(y)$ is a value if $g(\cdot)$ is one to one.
- If $g(\cdot)$ is many to one, then $g^{-1}(y) := \{x : g(x) = y\}$.
- In that case, $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in g^{-1}(y)\}$.
- $p_Y(y) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in \{x : g(x) = y\}\}$.
- Now $\mathbb{P}\{\omega \in \Omega : X(\omega) \in B\} = \sum_{\{x \in B\}} p_X(x)$ for $B \in \mathcal{F}'$.
- Proof follows after setting $B = \{x : g(x) = y\}$ $\square$

# $E[g(X)]$

Theorem: Suppose $Y = g(X)$ and $X$ is discrete with pmf $p_X(\cdot)$. Then, $E[Y] = \sum_x g(x)p_X(x)$

Proof

$$\begin{aligned} E[Y] &= \sum_y y p_Y(y) \\ &= \sum_y \sum_{\{x: g(x)=y\}} g(x) p_X(x) \\ &= \sum_x g(x) p_X(x). \end{aligned}$$

$\square$

https://en.wikipedia.org/wiki/Law_of_the_unconscious_statistician

# Towards Variance ..

- Recall $E[X] = \sum_{x \in \Omega'} x p_X(x)$.
- Furthermore, $E[X^n] = \sum_{x \in \Omega'} x^n p_X(x)$.
- In general, $E[g(X)] := \sum_{x \in \Omega'} g(x) p_X(x)$
- Now consider $g(X) = (X - E[X])^2$. $g(X)$ quantifies the square of the deviation of $X$ from the mean.
- Note $g(X)$ cannot track if the deviation is positive or negative!
- $E[g(X)]$ would then tell us the mean of the square of the deviation.
- In fact, $\sqrt{E(g(X))}$ quantifies the deviation.

# Variance

- $E[g(X)] = E[(X - E[X])^2]$ is called as the variance of random variable $X$.
- $Var(X) := E[(X - E[X])^2]$
- HW: Prove that $E[(X - E[X])^2] = E[X^2] - E[X]^2$
- $\sigma_X = \sqrt{Var(X)}$ is called as the standard deviation of $X$.
- For a fair coin toss, instead of $\Omega' = \{1, -1\}$, what if we use $\{+100, -100\}$ ? The latter has more variance!
- HW: What is $Var(Y)$ where $Y = aX + b$?

# Examples of discrete random variables

# Indicator random variable

- Indicator random variable $1_A(\omega) = \begin{cases} 1, & \text{If } \omega \in A \subseteq \Omega \\ 0, & \text{otherwise.} \end{cases}$
- Its PMF is $p_{1_A}(x) = \begin{cases} \mathbb{P}(A), & \text{when } x = 1 \\ 1 - \mathbb{P}(A), & \text{when } x = 0. \end{cases}$
- This is a discrete random variable even though $\Omega$ could be continuous.
- For example, Event $A$ could be that the number picked uniformly on the real line is positive.
- What is its CDF and mean denoted by $E[1_A]$?
- What about its mean variance and moments?

# Bernoulli random variable

- Bernoulli random variable $X = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{otherwise.} \end{cases}$
- This is same as an indicator variable but here we do not specify $A$.
- As a matter of convenience, we will start ignoring $\Omega$ from now on.
- These random variables are used in Binary classification in ML. $X = 1$ if image has a cat.
- Basic models of Multi-arm bandit problem assume Bernoulli Bandits.
- $E[X] = p, E[X^n] = p$.

# Binomial $B(n, p)$ random variable.

- Consider a biased coin (head with probability $p$) and toss it $n$ times.
- Denote head by 1 and tail by 0.
- Let random variable $N$ denote the number of heads in $n$ tosses.
- PMF of $N$?. $p_N(k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- HW: What is $E[N], E[N^2], Var(X)$?

# Geometric random variable

- Consider a biased coin (head with probability $p$) and suppose you keep tossing it till head appears the first time.
- Let random variable $N$ denote the number of tosses needed for head to appear first time.
- What is the PMF of $N$? $p_N(k) = (1-p)^{k-1}p$.
- HW: What is $E[N], E[N^2], Var(N)$?

# Poisson random variable

- A Poisson random variable $X$ comes with a parameter $\lambda$ and has $\Omega' = \mathbb{Z}_{\geq 0}$
- PMF: $p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$
- Intuitively its a limiting case of the Binomial distribution with $n$ increasing and $p$ decreasing such that $np$ converges to $\lambda$.
- Mean of binomial is $np$ so $p$ should decrease while $n$ increases.
- Read the Wiki page on Poisson limit theorem.
- We will see more of this when we see Poisson Prcoesses.