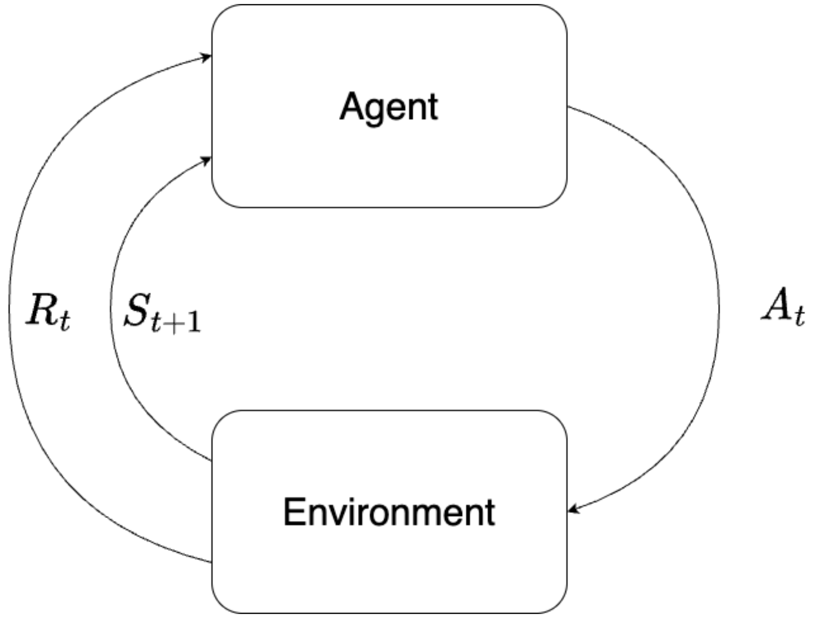
CMPUT 655 Lecture 2: Probability Theory Intro

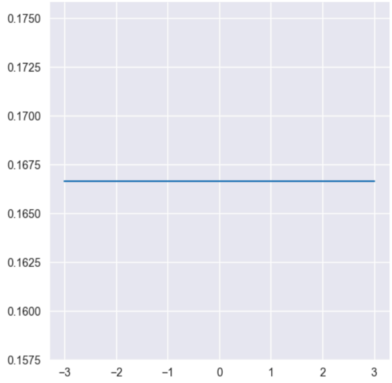
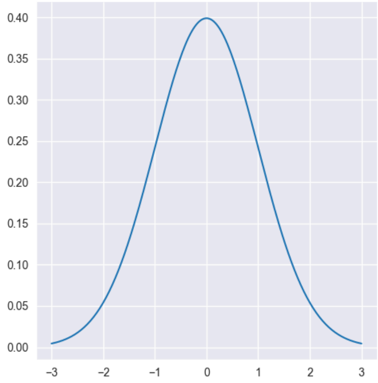
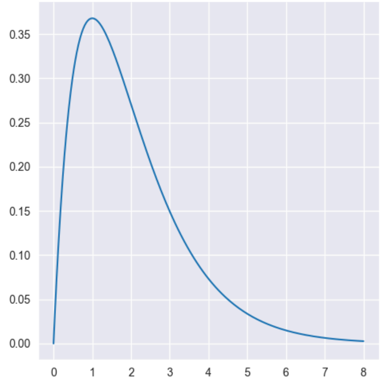
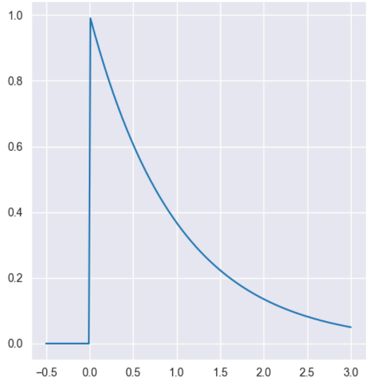
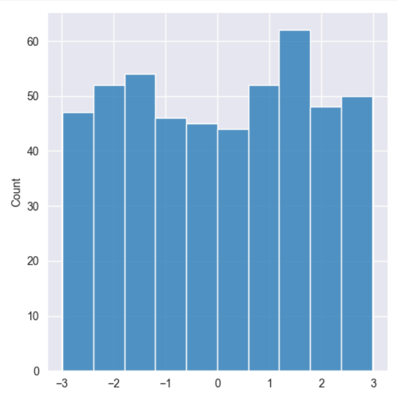
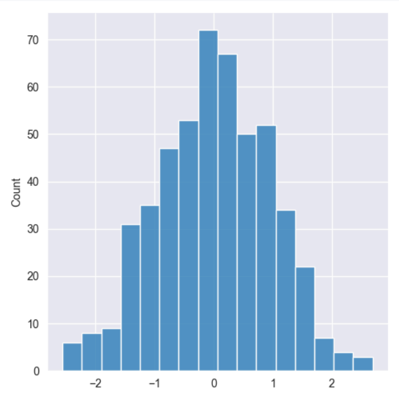
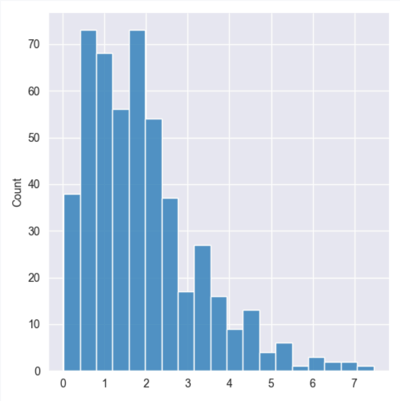
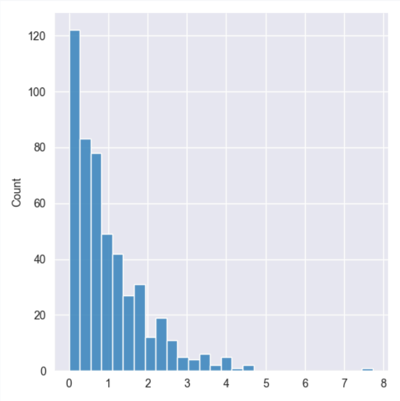
What is a random variable?

Random Variables: in Context

- R_t – reward in a given state following a given action
- S_{t+1} – next state, given a previous state and an action
- A_t – action chosen by our policy



Distributions

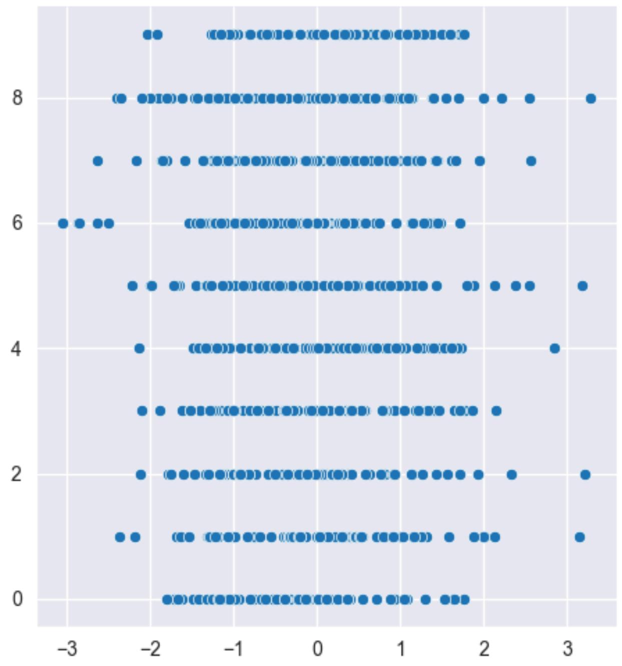
	Uniform	Gaussian	Gamma	Exponential
Density				
Samples				

Distributions: Notation

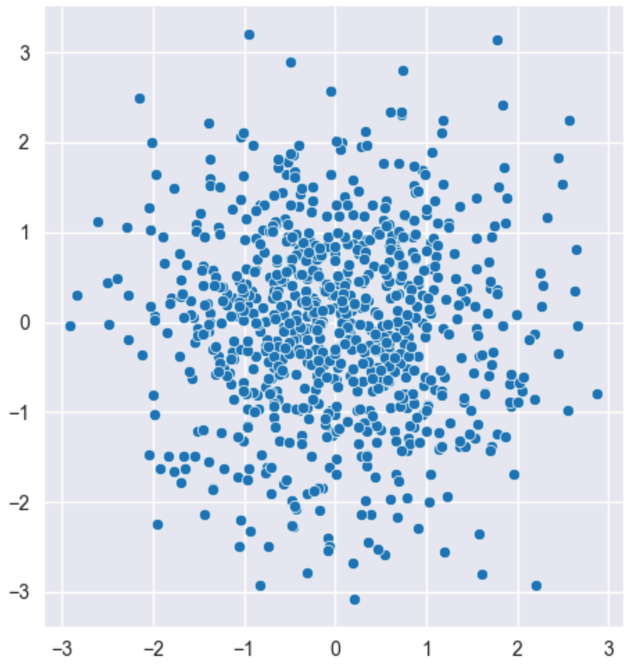
	Discrete	Continuous	
PDF	$\Pr\{\xi = x\} = p_\xi(x)$	$\Pr\{\xi = x\} \neq f_\xi(x)$	sums/integrates to 1
CDF	$F_\xi(x) = \Pr\{\xi \leq x\} = \sum_{\omega \leq x} p(\omega)$	$F_\xi(x) = \int_{-\infty}^{x} f_\xi(x) dx$	0 as $x \to -\infty$, and 1 as $x \to +\infty$

Joint Distributions

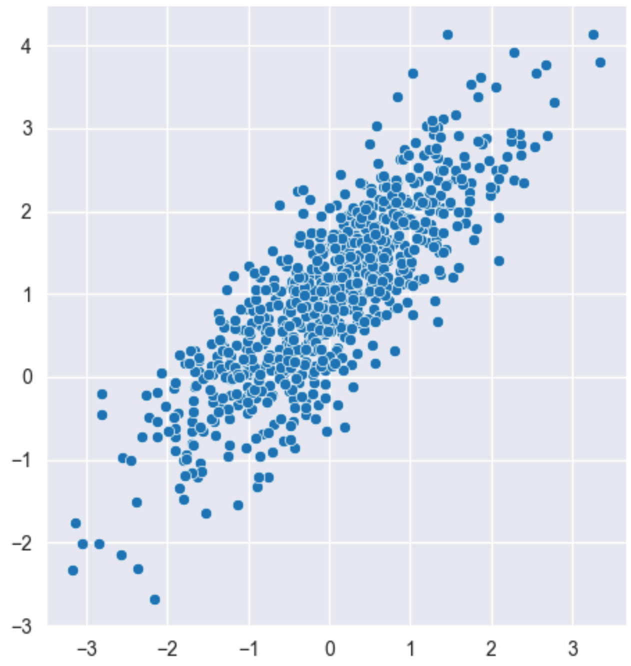
Gaussian + Categorical



Gaussian + Gaussian



Gaussian + Gaussian



Marginal Distributions

Discrete	Continuous
$\Pr\{\xi = x\} = \sum_{i=0}^{N} \Pr\{\xi = x, \eta = y_i\}$	$f_\xi(x) = \int_{-\infty}^{+\infty} f_{\xi,\eta}(x, y) dy$

Conditional Distributions

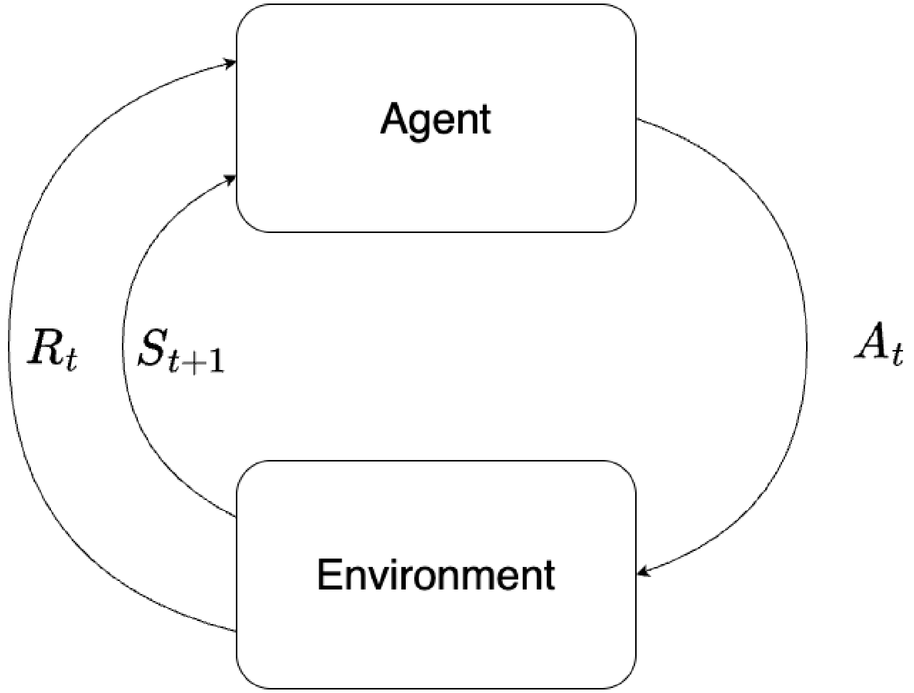
Discrete	Continuous
$\Pr\{\xi = x \mid \eta = y\} = \dfrac{\Pr\{\xi = x, \eta = y\}}{\Pr\{\xi = x\}}$	$f_\xi(x \mid y) = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$

Bayes Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Distributions: in Context

- $\Pr\{R_t = r \mid S_t, A_t\}$ or $f(r|S_t, A_t)$
- $\Pr\{S_{t+1} = s \mid S_t, A_t\}$
- $\pi(a|S_t)$



Independance

ξ and η are independant if:

- $F_{\xi,\eta}(x, y) = F_\xi(x) \cdot F_\eta(y)$
- $\Pr\{\xi = x | \eta = y\} = \Pr\{\xi = x\} \, \forall x, y$
- $f_\xi(x|y) = f_\xi(x) \, \forall x, y$

Expectation

Discrete	Continuous	Empirical Approx.
$\mathbb{E}\xi = \sum_{i=0}^{N} p(x_i) \cdot x_i$	$\mathbb{E}\xi = \int_{-\infty}^{\infty} x f_\xi(x) dx$	$\mathbb{E}\xi \approx \frac{1}{N} \sum_{i=0}^{N} \xi_i$

Mode & Median

	Discrete	Continuous
Mode	$\arg\max p(x_i)$	$\arg\max f_\xi(x)$
Median	$x_0 : \Pr\{\xi < x_0\} = \Pr\{\xi > x_0\} = 0.5$	$F_\xi(x_0) = 0.5$

Conditional Expectation

Discrete	Continuous
$\mathbb{E}[\xi \| \eta = y] = \sum_{i=0}^{N} x_i \mathrm{Pr}\{\xi = x_i \| \eta = y\}$	$\mathbb{E}[\xi \| \eta = y] = \int_{-\infty}^{+\infty} x f_{\xi}(x \| y) dx$

Example #1

(on p. 49 of the RL book)

$$
\begin{aligned}
r(s,a) &\doteq \mathbb{E}[R_t | S_{t-1} = s, A_{t-1} = a] \\
&= \sum_{r \in \mathcal{R}} r \cdot \Pr\{R_t = r | S_{t-1} = s, A_{t-1} = a\} \\
&= \sum_{r \in \mathcal{R}} r \underbrace{\sum_{s' \in \mathcal{S}} \Pr\{S_{t+1} = s', R_t = r | S_{t-1}, A_{t-1} = a\}}_{\text{marginalizing out the next state}}
\end{aligned}
$$

Example #2 (Bellman Equation)

(on p. 59 of the RL book)

$$
\begin{aligned}
v &\doteq \mathbb{E}_\pi[G_t | S_t = s] \\
&= \mathbb{E}_\pi[R_t + \gamma G_t | S_t = s] \\
&= ?
\end{aligned}
$$

Variance

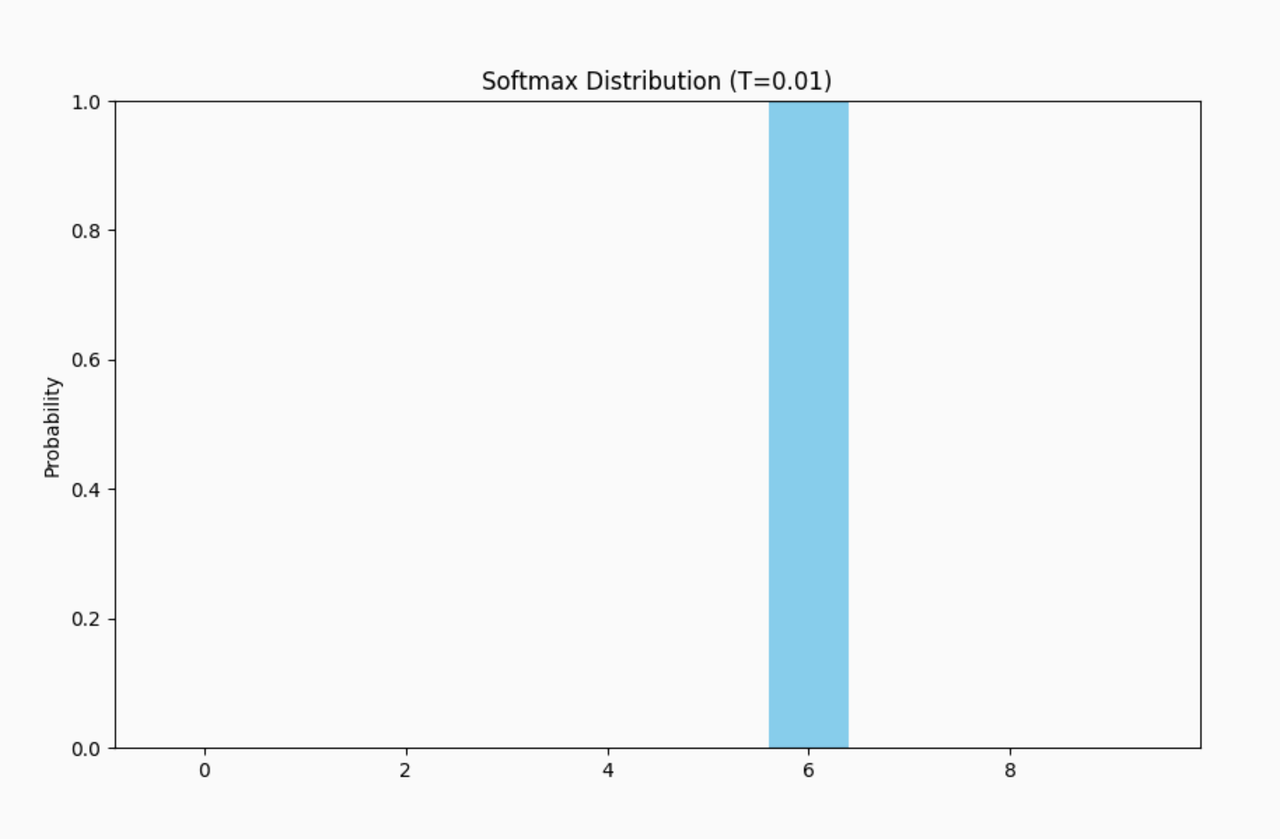
$$\mathrm{Var}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = \begin{cases} \displaystyle\sum_{i=0}^{N}(x_i - \mathbb{E}\xi)^2 \mathrm{Pr}\{\xi = x_i\}, \text{ discrete,} \\ \displaystyle\int_{-\infty}^{+\infty}(x - \mathbb{E}\xi)^2 f_\xi(x)dx, \text{ continuous} \end{cases}$$

Entropy

	Discrete (Shannon Entropy)	Continuous (Differential Entropy)
$H(\xi)$	$-\sum_{x \in \mathcal{X}} p(x) \log p(x)$	$-\int_{\mathcal{X}} f(x) \log f(x) dx$

Softmax / Gibbs / Boltzmann

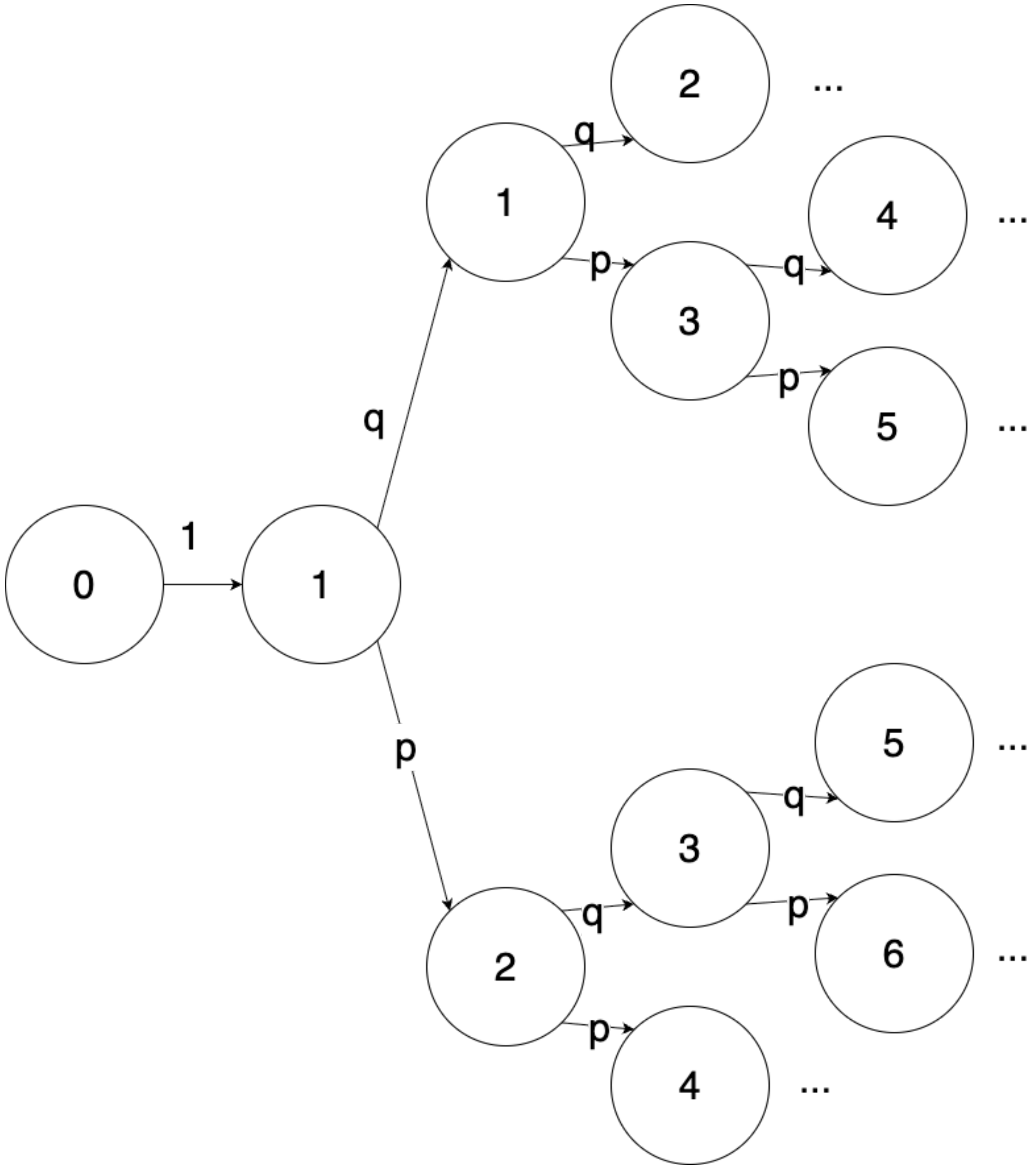
$$\text{softmax}(\vec{x};T)_i = \frac{e^{\frac{x_i}{T}}}{\sum_{j=0}^{N} e^{\frac{x_j}{T}}}$$



Markov Processes

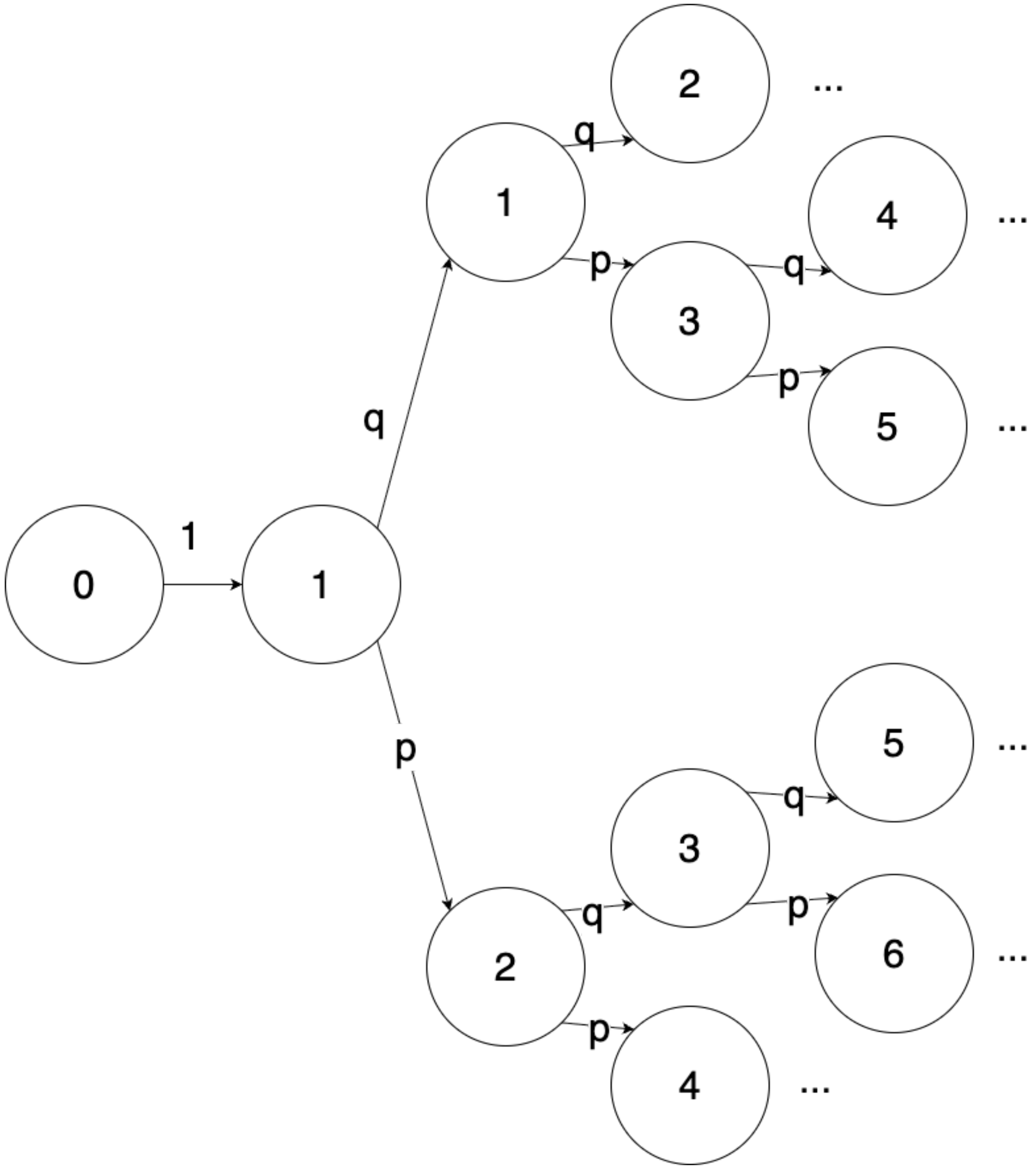
Markov Property. The future depends only on the present:

$$\Pr\{X_{t+1} | X_t, X_{t-1}, \ldots, X_0\} \doteq \Pr\{X_{t+1} | X_t\}$$



Example #3

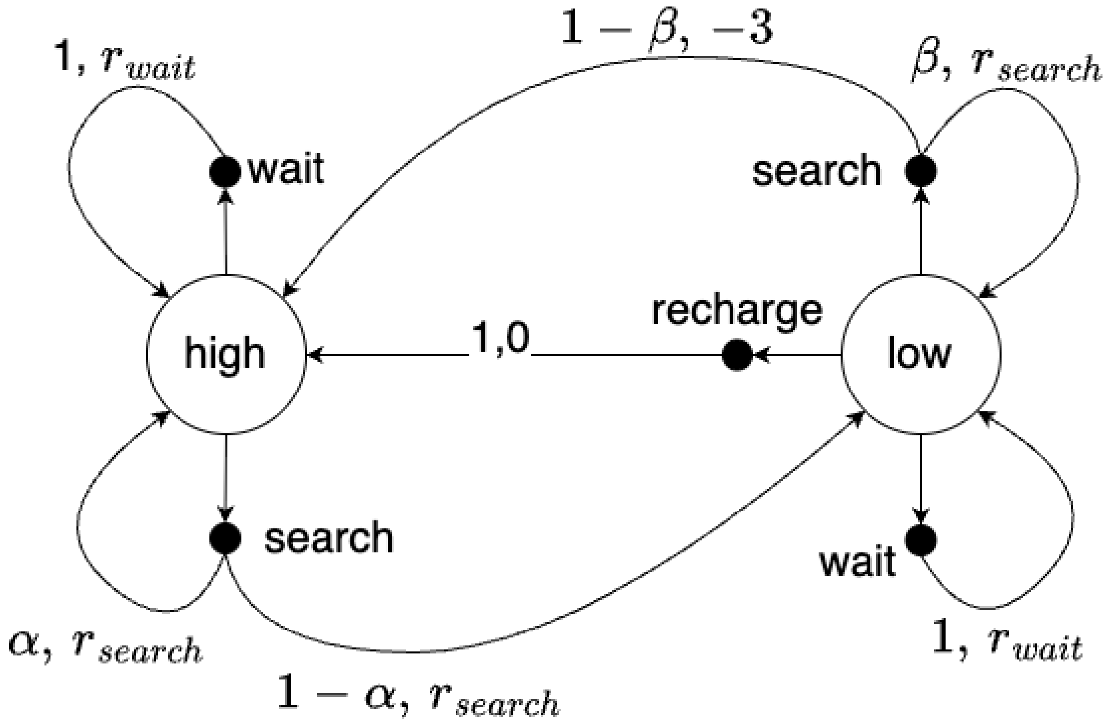
- $X_0 = 0,\ X_1 = 1$
- $X_n = X_{n-1} + X_{n-2} + \varepsilon_n$
- ε_n – iid, $\Pr\{\varepsilon_n = 1\} = p$, $\Pr\{\varepsilon_n = 0\} = q = 1 - p$



Markov Decision Processes

$(\mathcal{S}, \mathcal{A}, p, r)$:

- $S_{t+1} \sim p(s'|S_t = s, A_t = a)$
- $r(s, a) = \mathbb{E}[R_t|S_t = s, A_t = a]$



Q&A