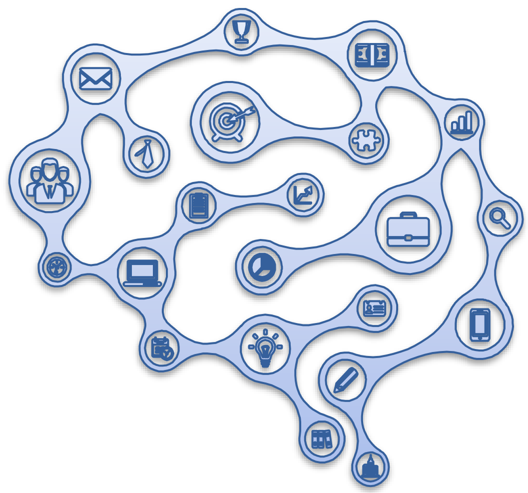
의료 Artificial Intelligence

확률과 인공지능



2022.04.14

오늘 배울 내용 …

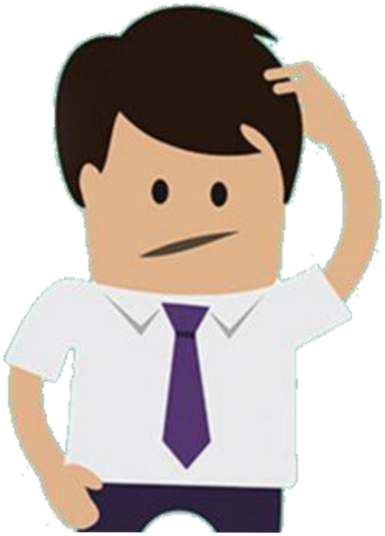
1. 확률과 통계

2. 논리설계 실습

3. mblock 실습

어렵지 않다

쉬운 것도 아니다



인공지능 이론

확률 통계 기초

167 161 178

183

178

171

161

171 165

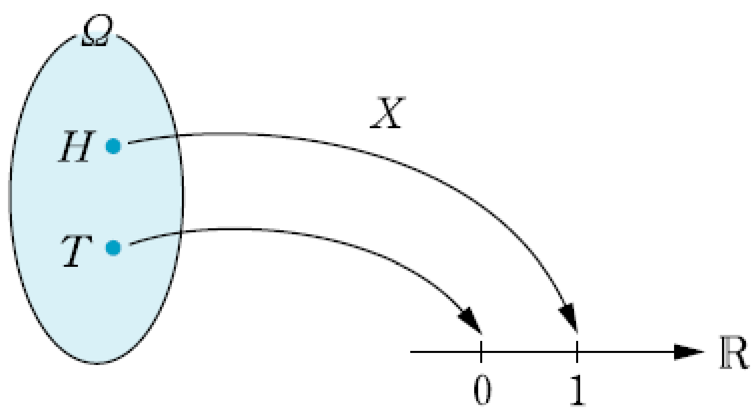
178

161 165 167 171 178 183

확률 통계 기초

확률변수

- 표본공간 Ω의 각 원소에 하나의 실숫값을 대응하는 함수 X를 **확률변수(random variable)**라고 한다.



[그림 3-1] 확률변수의 정의

확률 통계 기초

이산확률변수

확률변수 X의 치역이 셀 수 있는 이산값으로 주어지는 확률변수 X를

이산확률변수(discrete random variable)라 한다.

$$X(\Omega) = \{1, 2, 3, 4, \cdots\}$$

연속확률변수

- 어떤 연속하는 범위 안에서 모든 실숫값을 가지는 확률변수 X를

  연속확률변수(continuous random variable)라 한다.

$$X(\Omega) = \{x \in \mathbb{R},\ 0 \le x \le 100\}$$

확률 통계 기초

확률질량함수

- 이산확률변수 X에 대하여 X가 임의의 실수 x를 취할 확률에 대응하는 다음 함수를 이산확률변수 X의 **확률질량함수(probability mass function)**라 한다.

$$f(x) = \mathrm{P}(X = x)$$

확률질량함수의 성질

- 이산확률변수 X의 확률질량함수 $f(x)$에 대하여 다음 성질이 성립한다.

(1) 모든 실수 x에 대하여 $0 \le f(x) \le 1$이다.

(2) $\sum_{x \in \mathbb{R}} f(x) = 1$이다.

(3) 임의의 $A \subset \mathbb{R}$에 대하여 $\mathrm{P}(X \in A) = \sum_{x \in A} f(x)$이다.

확률 통계 기초

확률밀도함수

- 연속확률변수 X에 대하여 $a \le X \le b$일 확률을 다음과 같이 표현할 때, 확률변수 X는 연속확률분포를 따른다.

$$P(a \le X \le b) = \int_a^b f(x)dx$$

- 이때 연속함수 $f(x)$를 확률변수 X의 **확률밀도함수(probability density function)**라 한다.

확률밀도함수의 성질

- 연속확률변수 X의 확률밀도함수 $f(x)$에 대하여 다음 성질이 성립한다.

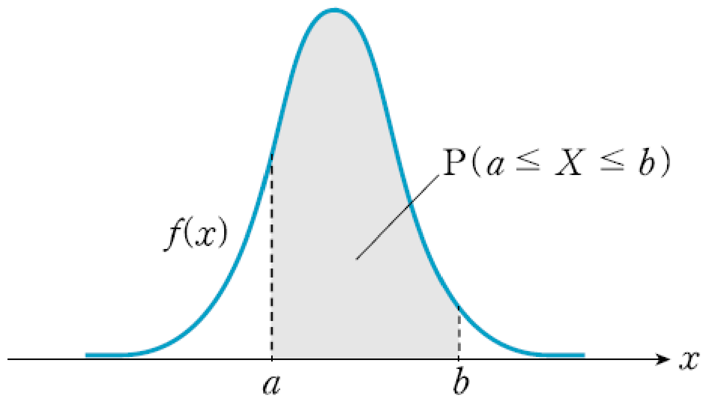
(1) 모든 실수 x에 대하여 $f(x) \ge 0$이다.

(2) $\int_{-\infty}^{\infty} f(x)dx = 1$이다.

(3) $P(a \le X \le b) = \int_a^b f(x)\,dx$이다.

확률 통계 기초

확률변수 X가 a와 b 사이에 있을 확률 $\mathrm{P}(a \le X \le b)$는 $x = a$와 $x = b$, 그리고 연속함수 $f(x)$의 그래프와 x축으로 둘러싸인 면적과 같다. 따라서 연속확률변수 X에 대하여 다음이 성립함을 알 수 있다.



[그림 3-3] 연속확률변수 X의 확률 $\mathrm{P}(a \le X \le b)$

- $\mathrm{P}(X = a) = \int_a^a f(x)dx = 0$
- $\mathrm{P}(a \le x \le b) = \mathrm{P}(a \le x < b) = \mathrm{P}(a < x \le b) = \mathrm{P}(a < x < b)$

확률 통계 기초

다음과 같은 함수 $F(x)$ 를 확률변수 X 의 **분포함수(distribution function)** 라 한다.

$$F(x) = \mathrm{P}(X \leq x)$$

누적분포함수

- 이산확률변수 X의 분포함수는 임의의 실수 x에 대하여 이산확률변수 X가 x 보다 작거나 같은 값을 취하는 확률로 정의로 정의하므로, 분포함수 $F(x)$를 **누적분포함수(cumulative distribution function)** 라고도 한다.

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

- $f(x)$는 확률질량함수이다.

확률 통계 기초

분포함수의 성질

- 분포함수 $F(x)$ 에 대하여 다음 성질이 성립한다.

(1) 모든 실수 x 에 대하여 $0 \le F(x) \le 1$ 이다.

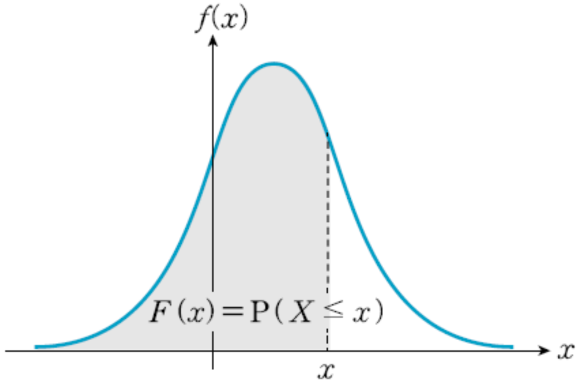
(2) $F(x)$ 는 증가함수이다.

(3) $F(\infty) = \lim_{x \to \infty} F(x) = 1$, $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ 이다.

(4) X 가 이산확률변수인 경우, $\mathrm{P}(X = x) = F(x) - F(x-1)$ 이다.

$f(x)$가 연속확률변수 X의 확률밀도함수일 때, 확률변수 X의 분포함수 $F(x)$는 다음과 같다.

$$F(x) = \mathrm{P}(X \le x) = \int_{-\infty}^{x} f(x)\,dx$$



- 분포함수를 이용하여 연속확률변수 X가 구간에 있을 확률을 구할 수 있다.

$$\mathrm{P}(a \le X \le b) = \mathrm{P}(X \le b) - \mathrm{P}(X \le a) = F(b) - F(a) = \int_{a}^{b} f(x)\,dx$$

확률 통계 기초

통계에서 2개의 주요한 값 : 평균(=값의 특징), 분산 (=분포의 특징)

평균 → '보통 얼마야…' 하는 값

데이터 집단을 대표하는 값

데이터 집단에서 임의로 데이터를 뽑을 때 확률적으로 가장 많이 나오는 값

수집된 데이터 $X_1, X_2, \cdots, X_N$에 대해서 **평균**은 다음과 같으며, 기호는 $E(X)$로 표기합니다.

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} X_i$$

확률 통계 기초

변수에 대해 확률이 균일하지 않을 때(각각 확률이 있을 때) 평균

확률변수 X의 기댓값

- 확률변수 X의 **기댓값(expected value)** $\mathrm{E}(X)$는 다음과 같이 정의한다.

$$\mathrm{E}(X) = \begin{cases} \displaystyle\sum_{i=1}^{n} x_i f(x_i) & (X: \text{이산확률변수}) \\ \displaystyle\int_{-\infty}^{\infty} x f(x) dx & (X: \text{연속확률변수}) \end{cases}$$

확률 통계 기초

분산 → 데이터 간의 거리가 어떤 가? (중심에서 얼마나 떨어져 있는 가)

임의로 뽑은 값이 평균과의 차이가 어떤 가?

- 평균으로 예측했을 때 얼마나 믿을 수 있는 지

- 데이터의 분포 정도가 어떤 지

수집된 데이터 $X_1, X_2, \cdots, X_N$의 평균을 $E(X)$라 할 때 수집된 데이터의 **분산**과 **표준편차**는 다음과 같으며, 기호로는 각각 $V(X)$, $\sigma(X)$로 표기합니다.

$$V(X)=\frac{1}{N}\sum_{i=1}^{N}(X_i-E(X))^2$$

$$\sigma(X)=\sqrt{V(X)}$$

$$V(X)=E(X^2)-[E(X)]^2$$

확률 통계 기초

변수에 대해 확률이 균일하지 않을 때(각각 확률이 있을 때) 분산

- $\mu = \mathrm{E}(X)$ 라 할 때, 확률변수 X 의 **분산(variance)**은 다음과 같이 정의된다.

$$\mathrm{Var}(X) = \mathrm{E}\{(X-\mu)^2\} = \begin{cases} \displaystyle\sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) & (X : \text{이산확률변수}) \\ \displaystyle\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx & (X : \text{연속확률변수}) \end{cases}$$

그리고 X 의 표준편차는 σ_X 로 표기하여 다음과 같이 정의한다.

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

확률 통계 기초

5일 동안 어느 지역의 일평균 기온에 대한 데이터를 수집했습니다. 다음 데이터를 바탕으로 일평균 기온의 분산과 표준편차를 구하세요.

	1일	2일	3일	4일	5일
기온(℃)	20	24	22	18	21

확률 통계 기초

어떤 프로야구선수가 한 게임에서 치는 안타 수와 그 확률을 조사하였더니 다음 표와 같았다.
이 선수가 한 게임에서 치는 평균 안타 수를 구하라. 그리고 분산과 표준편차를 구하라.

X	0	1	2	3	4	합계
$\mathrm{P}(X=x)$	0.30	0.35	0.20	0.10	0.05	1

$$\mu = \mathrm{E}(X) = 0\times 0.30 + 1\times 0.35 + 2\times 0.20 + 3\times 0.10 + 4\times 0.05$$

$$= 1.25$$

$$\mathrm{Var}(X) = \mathrm{E}\{(X-\mu)^2\} = \sum_{i=1}^{5}(x_i-\mu)^2 f(x_i)$$

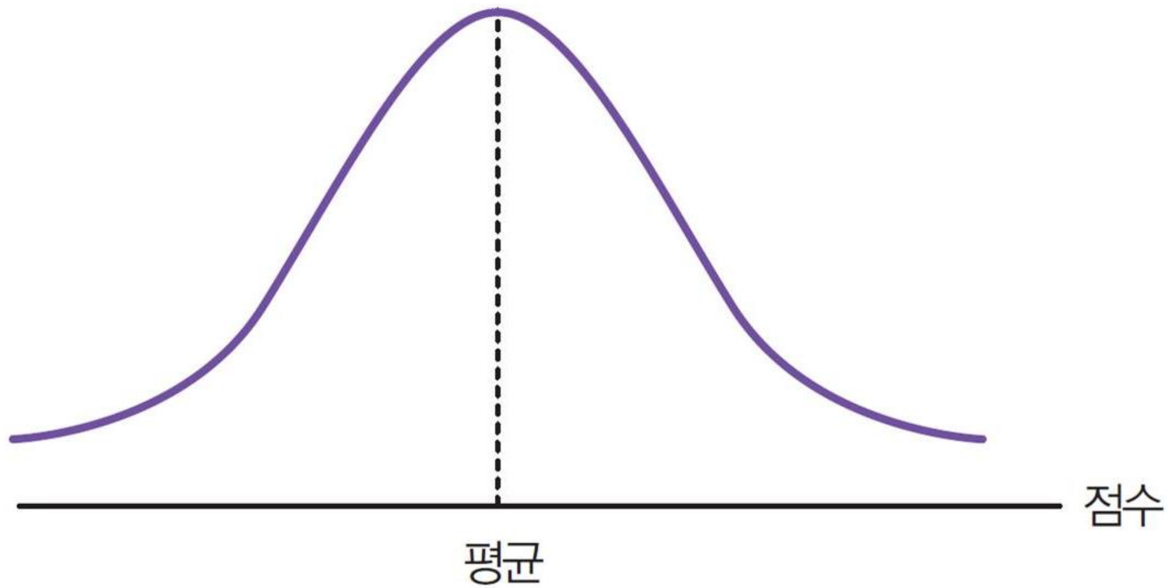
$$=(0-1.25)^2\times 0.30 + (1-1.25)^2\times 0.35 + (2-1.25)^2\times 0.20$$

$$+(3-1.25)^2\times 0.10 + (4-1.25)^2\times 0.05 = 1.2875$$

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{1.2875} \fallingdotseq 1.135$$

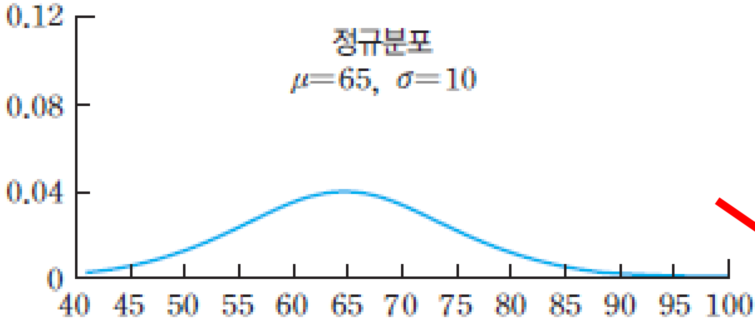
확률 통계 기초

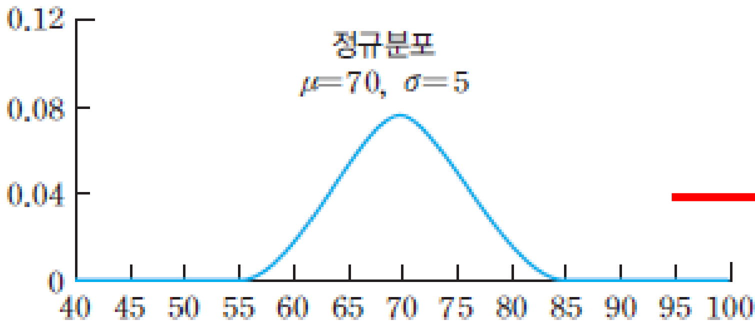
- Q : 현실에서의 여러가지 변수들의 분포를 잘 설명하는 확률밀도함수는 없을까?
- A : 정규분포!

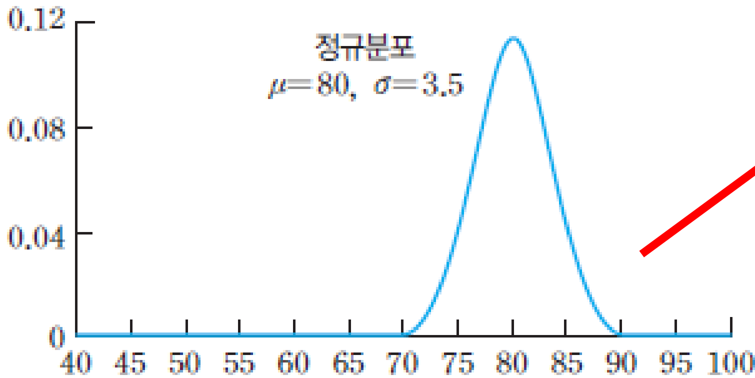


시험 점수의 분포

확률 통계 기초







표준화

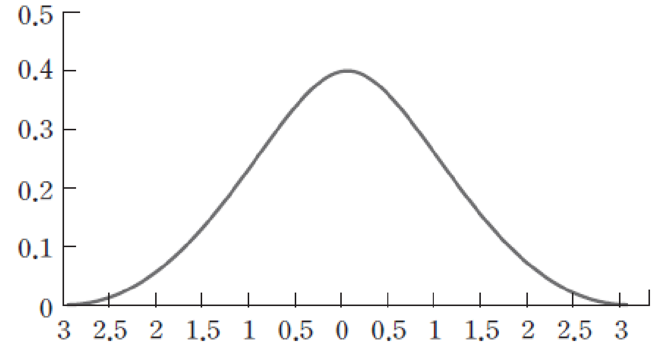
$$Z=\frac{X-\mu}{\sigma}$$

표준화

$$Z=\frac{X-\mu}{\sigma}$$

표준화

$$Z=\frac{X-\mu}{\sigma}$$



표준정규분포

$\mu=0,\ \sigma=1$

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$$

$$Z\sim N(0,\ 1)$$

확률 통계 기초

* 평균 μ와 표준편차 σ 가 다른 정규분포의 확률변수 X에 대해 **표준화를 하면**
이 분포는 **평균 μ=0, 표준편차 σ=1인 표준정규분포 N(0,1)**가 된다.

이때 표준화 식 $Z = \dfrac{X-\mu}{\sigma}$ 을 통해 얻어진 값을 Z점수(Z score)라 한다.

* **표준화** $Z = \frac{X-m}{\sigma}$: 평균을 0, 표준편차를 1로 변환

$$\mathrm{E}(Z) = \mathrm{E}\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma}\mathrm{E}(X) - \frac{m}{\sigma} = \frac{m}{\sigma} - \frac{m}{\sigma} = 0$$

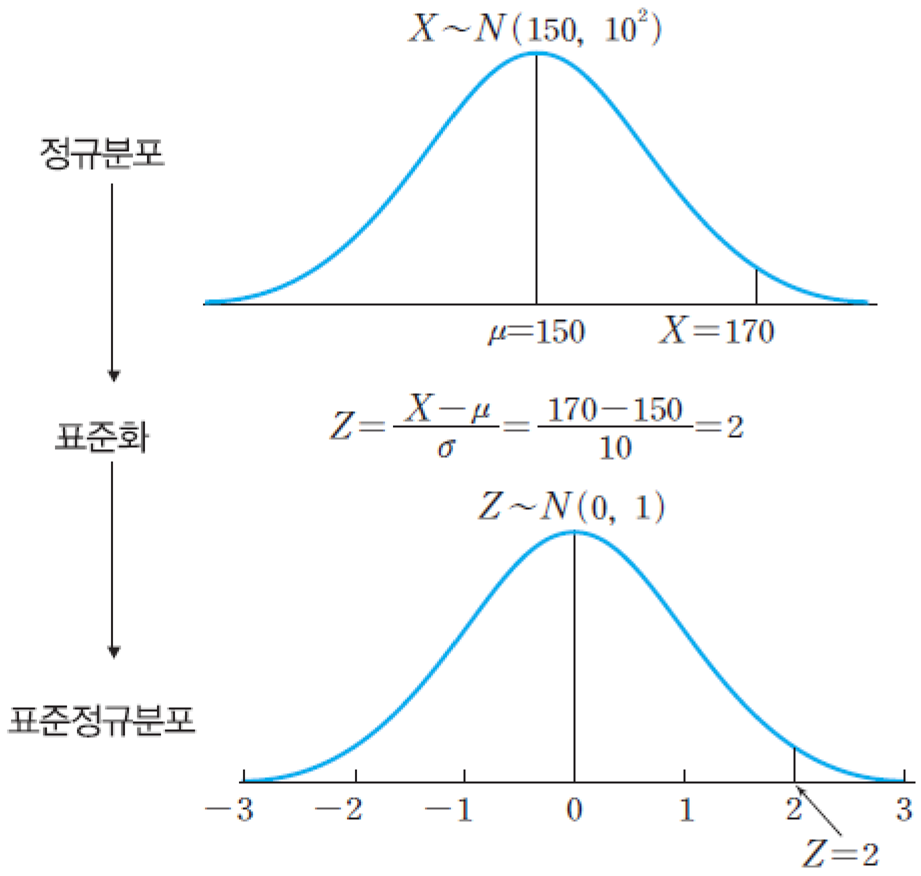
$$\mathrm{V}(Z) = \mathrm{V}\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma^2}\mathrm{V}(X) = \frac{1}{\sigma^2} \times \sigma^2 = 1$$

확률 통계 기초

평균 μ=150 이고, 표준편차 σ=10 인 정규분포 N(150, 10^2)에 대해 X=170에 대한 Z 점수를 계산하시오.

표준화 식 $Z=\frac{X-\mu}{\sigma}$ 을 이용하면

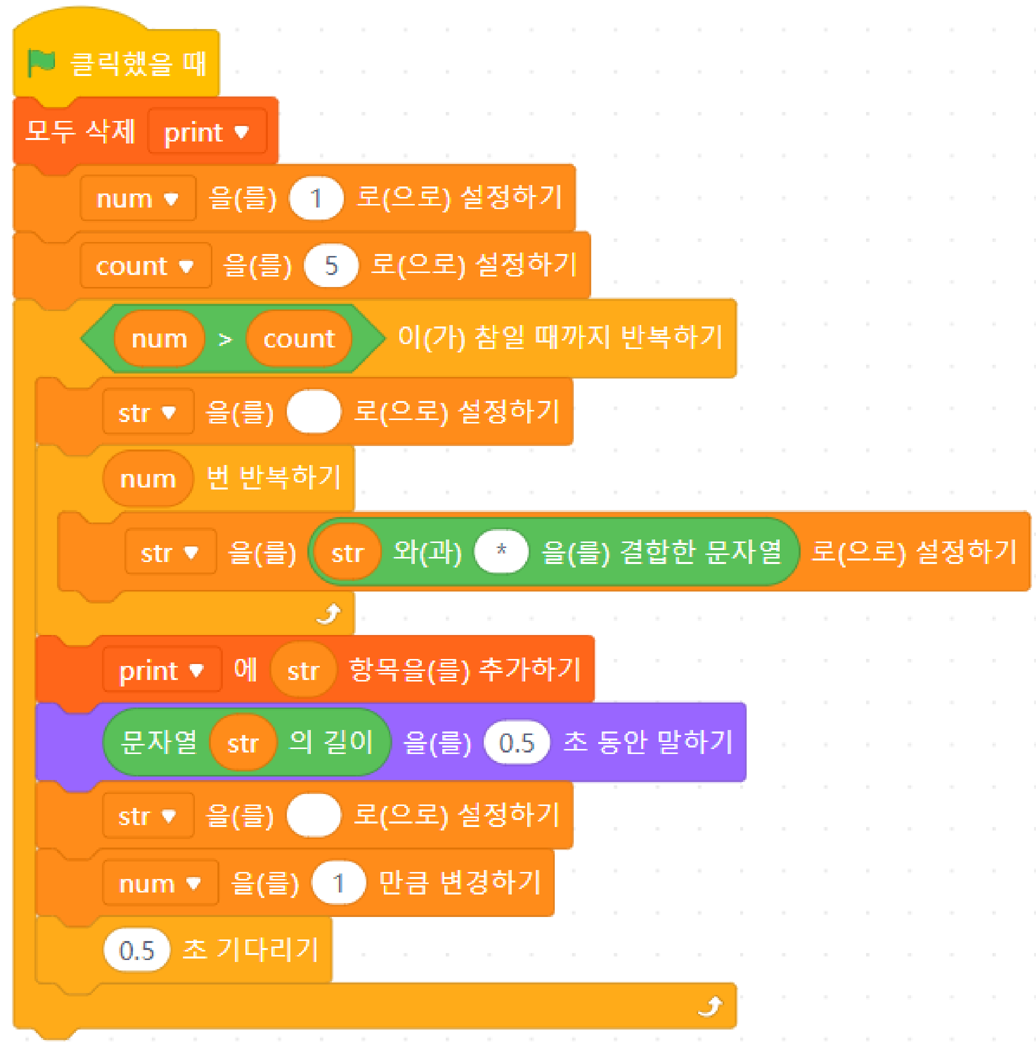
$Z=\frac{170-150}{10}$ 이므로 점수는 2.



지능 만들기 - 논리설계 실습

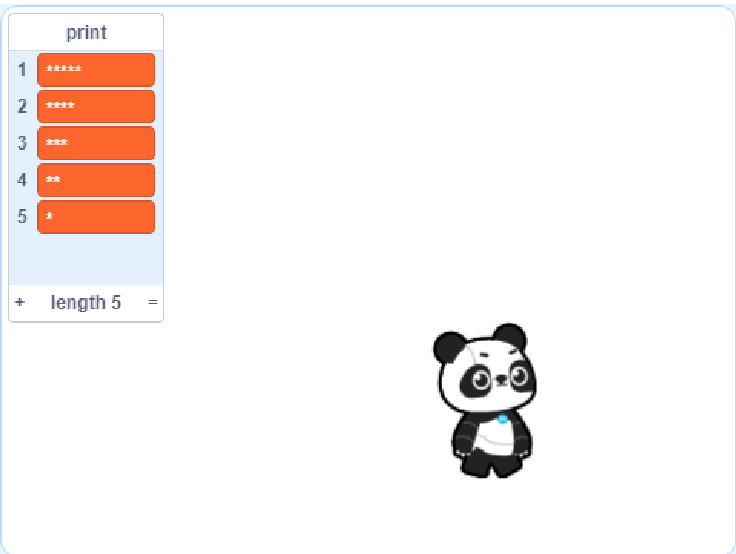
논리설계 실습 - 1

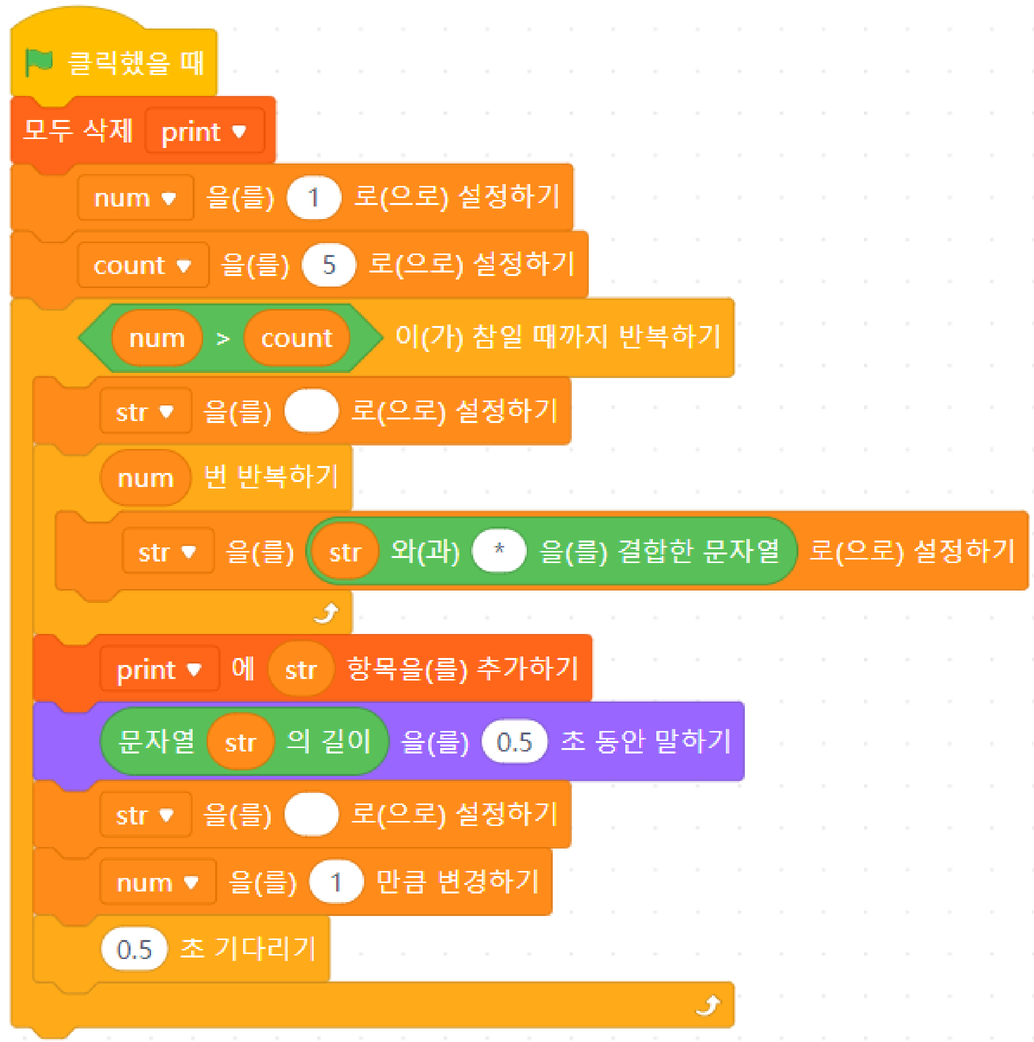
다음 블록프로그램 수행되면
list에 담기는 데이터를 예상해 보시오.



논리설계 실습 - 2

실습 1의 리스트 값의 순서가 반대로 저장 되도록 바꾸기

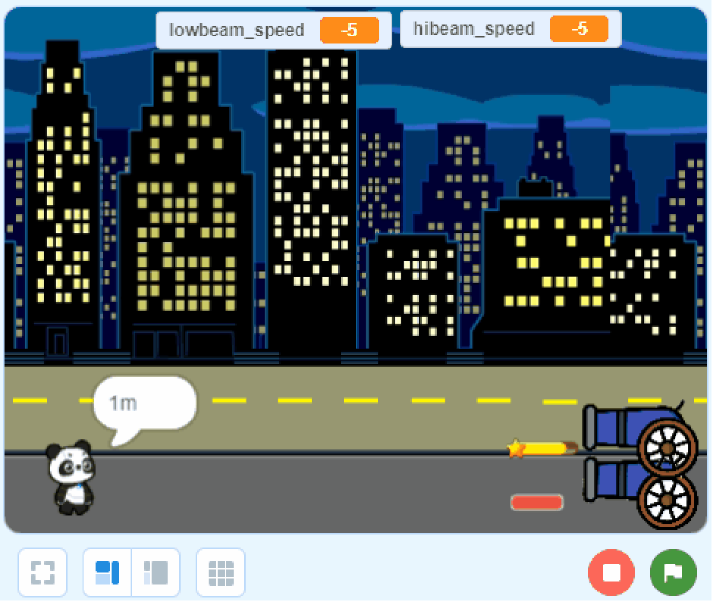




MBlock 실습

게임 만들기

총알 피하기 게임



1단계 : 분해

1. 펜더 객체
- 달리기 모션
- 점프 : 1단, 2단
- 숙이기

2. 하단/상단 포탄 객체
- 포탄 생성
- 좌로 이동
- 모양 변함

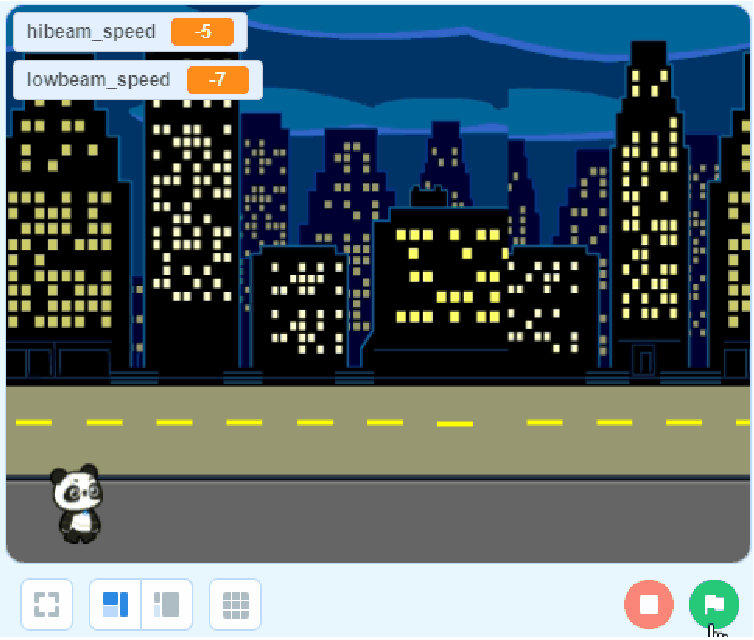
3. 배경
- 횡스트롤

4. 대포 객체
- 거리가 30미터이면 모양 변경

5. 게임 룰
- 펜더가 포탄에 닿으면 수명이 줄어 듬
- 수명이 3개 이상 줄어 들면 게임 끝
- 펜더는 달리면서 이동 거리를 계산해서 말함
- 게임이 끝나면 'Game Over'
  게임 클리어시에는 'You win the Game' 출력

게임 만들기

총알 피하기 게임



2단계 : 추상화

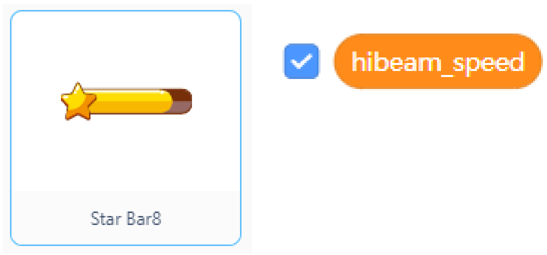
2. 포탄 객체

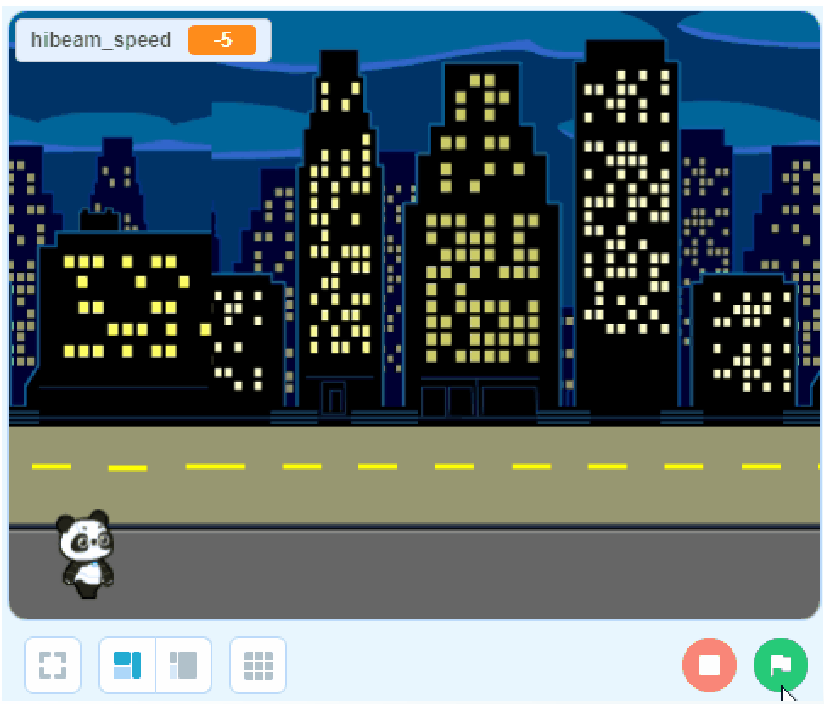
- 상단 포탄
  . 숨기기 / 보이기
  . 복제하기 / 복제본 삭제하기
  . 가장자리에 닿을 때까지 반복하기
  . x좌표 변경하기
  . 다음 모양 바꾸기
  . 기다리기 (복제본 만들기 인터벌)

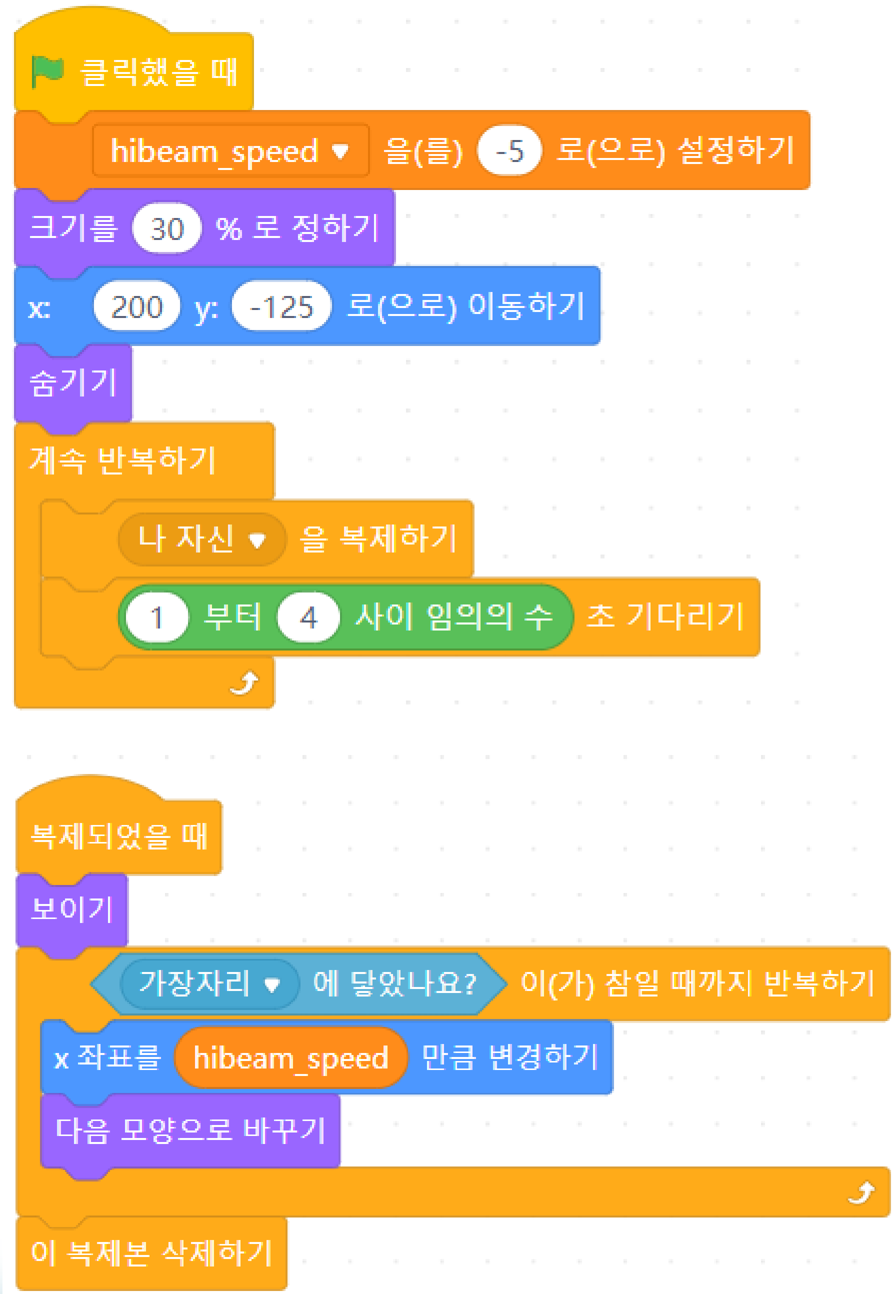
- 하단 포탄
  . 숨기기 / 보이기
  . 복제하기 / 복제본 삭제하기
  . 가장자리에 닿을 때까지 반복하기
  . x좌표 변경하기
  . 다음 모양 바꾸기
  . 기다리기 (복제본 만들기 인터벌)

게임 만들기 - 1

- 상단 포탄 추가

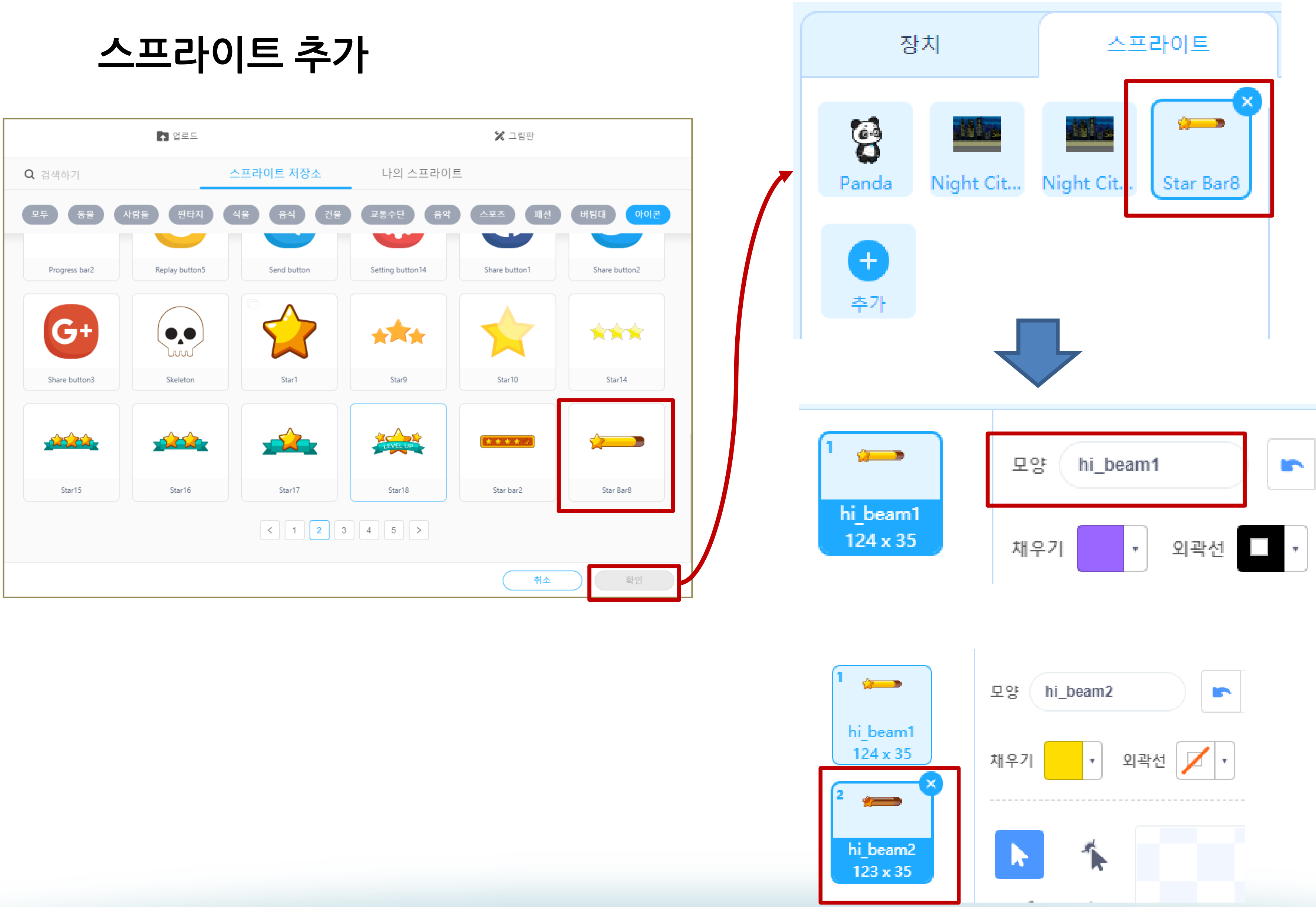






게임 만들기 - 2

스프라이트 추가



게임 만들기 - 2

- 하단 포탄 추가