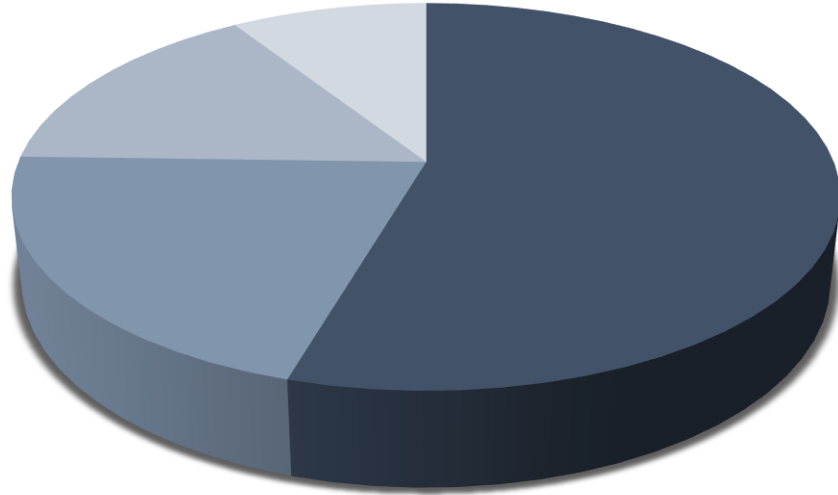




통계학 실습

10. 확률분포 (1)



- I. Binomial Distribution
- II. Geometric Distribution

Binomial Distribution

- 성공 확률이 P 인 베르누이 시행을 N 번 반복했을 때, 그 중에 X 번 성공할 확률은 Binomial Distribution을 따른다.
- 베르누이 시행
 - 어떤 사건에 대하여 사건의 결과가 ‘성공’ 혹은 ‘실패’로 나뉘는 경우를 일컫는다.

Binomial Distribution 내장 함수

함수 사용	설명	
<code>dbinom(x, size, prob)</code>	'D'ensity	확률 밀도 함수 결과값 구하기 ex) $P[X=x] = ?$
<code>pbinom(q, size, prob, lower.tail)</code>	'P'robability	누적 분포 함수의 누적확률 구하기 ex) $P[X \leq q] = ?$
<code>qbinom(p, size, prob, lower.tail)</code>	'Q'uantile	누적 확률에 해당하는 분위값 구하기 ex) $P[X \leq ?] = p$
<code>rbinom(n, size, prob)</code>	'R'andom	분포 함수를 따르는 난수 생성

- `x, q` 분위수 벡터
- `p` 확률 벡터
- `n` 추출 난수 개수

- `size` 시행 횟수
- `prob` 성공 확률
- `lower.tail` TRUE: $P[X \leq x]$, FALSE: $P[X > x]$

Binomial Distribution 문제 해결

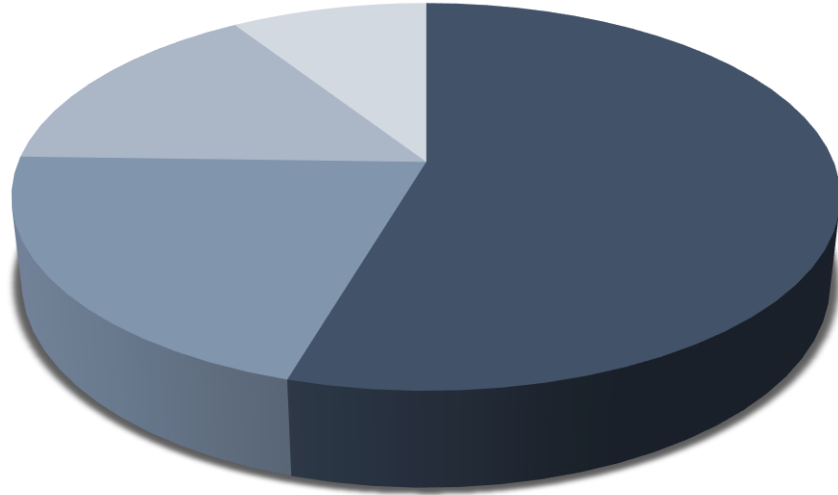
- 친구와 동전 내기를 하였다. 동전을 20번 던져서 앞면이 뒷면보다 더 많이 나올 경우 나의 승리다. 앞면이 11번 나와서 내가 이길 확률은?
 - I. 동전의 확률은 50%이므로 0.5이다. ($\text{prob} = 0.5$)
 - I. 11번의 성공이므로 구하고자 하는 식은 $P[X=11]$ 이다. ($x = 11$)
 - II. 확률 밀도 함수 값을 구해야 하므로 ‘`dbinom`’ 함수를 이용한다.

Binomial Distribution 함수 사용

- **Q1. 동전을 20번 던져서 그 중 11번 앞면이 나올 확률은?**
 - `dbinom(x=11, size=20, prob=0.5)`
- **Q2. 동전을 20번 던져서 앞면이 최대 7번이 나올 확률은?**
 - `pbinom(q=7, size=20, prob=0.5, lower.tail=TRUE)`
- **Q3. 동전을 20번 던졌을 때, 해당 분포를 따르는 난수를 10개 추출하여라.**
 - `rbinom(n=10, size=1, prob=0.5)`
 - Random 함수에서 size는 시행 결과의 범위를 의미한다.
 - Example : 동전 2 가지 → 0:실패, 1:성공

실습 문제

- 한 축구 선수가 5번의 슈트를 쏘면 2개가 골로 연결된다. 이 선수가 속한 팀이 현재 2점 뒤쳐지고 있다. 이기려면 3점이 필요한데, 이 선수가 슈트를 쏠 기회는 10번 뿐이다. 이 선수로 교체할 경우 이 팀이 승리할 확률은?
 - 팀이 승리하려면 적어도 3점이 필요하다.
 - 따라서 구하고자 하는 확률은 $P[X > 2]$ 으로 구할 수 있다.



- I. Binomial Distribution
- II. Geometric Distribution

Geometric Distribution

- 성공 확률이 P 인 베르누이 시행을 처음 성공할 때까지 반복한 시행 횟수가 X 일 확률은 Geometric Distribution을 따른다.
- 성공 전까지 반복한 시행 횟수 X 는 총 실패 횟수가 된다.
- 베르누이 시행
 - 어떤 사건에 대하여 사건의 결과가 ‘성공’ 혹은 ‘실패’로 나뉘는 경우를 일컫는다.

Geometric Distribution 내장 함수

함수 사용	설명	
<code>dgeom(x, prob)</code>	'D'ensity	확률 밀도 함수 결과값 구하기 ex) $P[X=x] = ?$
<code>pgeom(q, prob, lower.tail)</code>	'P'robability	누적 분포 함수의 누적확률 구하기 ex) $P[X \leq q] = ?$
<code>qgeom(p, prob, lower.tail)</code>	'Q'uantile	누적 확률에 해당하는 분위값 구하기 ex) $P[X \leq ?] = p$
<code>rgeom(n, prob)</code>	'R'andom	분포 함수를 따르는 난수 생성

- `x, q` 분위수 벡터
- `p` 확률 벡터
- `n` 추출 난수 개수

- `prob` 성공 확률
- `lower.tail` TRUE: $P[X \leq x]$, FALSE: $P[X > x]$

Geometric Distribution 문제 해결

- 어느 농구 선수가 3점 슈트를 던졌을 때, 들어갈 확률이 55%라고 한다. 이 선수가 4번의 시도 끝에 3점 슈트를 성공하게 될 확률은?
 - I. 들어갈 확률이 55%이므로, 0.55이다. ($\text{prob} = 0.55$)
 - II. 4번의 시도이므로 구하고자 하는 식은 $P[X=4]$ 이다. ($x = 4-1$)
 - III. 확률 밀도 함수 값을 구해야 하므로 'dgeom' 함수를 이용한다.

- Q1. 어느 농구 선수의 3점 슛의 확률이 0.55일 때, 4번째에 첫 성공할 확률은?
 - `dgeom(x=3, prob=0.55)`
- Q2. 위의 농구 선수가 최대 3번의 시도 안에 슛을 성공할 확률은?
 - `pgeom(q=2, prob=0.55, lower.tail=TRUE)`
- Q3. 위의 농구 선수가 슛 성공 확률이 누적 99%가 되는 최대 시도 횟수는?
 - `qgeom(p=0.99, prob=0.55)`
 - 결과를 `pgeom`의 `q` 벡터에 대입하면, 서로 역이 됨을 알 수 있다.

실습 문제

- Q. 어떤 공장에서 물건들의 결함을 조사하고 있다. 결함이 있으면 폐기, 결함이 없으면 유통을 한다. 이 때, 결함이 있을 확률이 P 이다. 100개의 물건 중 15개 이하가 결함일 확률이 0.2보다 작게 되는 결함의 확률 P 를 구하여라.
- HINT) 확률 P 값은 0부터 0.001씩 증가하여 조사한다.