

Sampling with and without replacement

Quiz section of a stats class, with 38 students

4 Freshman, 12 sophomores, 13 Juniors, 8 Seniors, 1 non-matriculated.

Q If we sample 2 students, with replacement, what is the prob that they are the same class level?

$$P(2 \text{ same class level}) = P(2 \text{ Fresh}) + P(2 \text{ soph}) + P(2 \text{ Jr}) + P(2 \text{ Sr}) + P(2 \text{ nm})$$

$$= \frac{4}{38} \times \frac{4}{38} + \frac{12}{38} \times \frac{12}{38} + \frac{13}{38} \times \frac{13}{38} + \frac{8}{38} \times \frac{8}{38} + \frac{1}{38} \times \frac{1}{38}$$

$$\approx 0.2729 = 27.29\%$$

Q If we sample five students, with replacement, what is the prob that 3 (exactly) are Juniors? 13 Juniors, 25 non-Juniors

$$P(3 \text{ Juniors out of } 5) = \frac{P(\text{Junior})P(\text{Junior})P(\text{Junior})P(\text{non-Junior})P(\text{non-Junior}) \times \text{Something}}{P(\text{JJJNN})}$$

$$\rightarrow \left(\frac{5}{3}\right) = \frac{5!}{3!2!}$$

Binomial Coefficient

There are multiple orders in which I can draw 3 Jrs and 2 non-Jrs

$$\binom{n}{k} \rightarrow n \text{ choose } k = \frac{n!}{k!(n-k)!}$$

$$n! = 1 \times 2 \times 3 \times \dots \times n$$

$$\begin{aligned} 5 \text{ items choose } 3 &= \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 2 \times 1} \\ &= \frac{5 \times 4}{2} = 10 \end{aligned}$$

List all combinations:

JJJNN	JJNJJ	JNNJJ
JJNJJ	JNJJJ	NNJJJ
JNJJJ	NNJJJ	NNJJJ
NNJJJ		

$$P(3 \text{ Juniors out of } 5) = \frac{13}{38} \times \frac{13}{38} \times \frac{13}{38} \times \frac{25}{38} \times \frac{25}{38} \times 10 = 17.33\%$$

Q What happens when we sample without replacement?

If we select 2 students, what is prob they are from the same class?

$$P(2 \text{ students from same class}) = P(2 \text{ Fresh}) + P(2 \text{ soph}) + P(2 \text{ Jr}) + P(2 \text{ Sr}) + P(2 \text{ non-Mat})$$

$$= \frac{4}{38} \times \frac{3}{37} + \frac{12}{38} \times \frac{11}{37} + \frac{13}{38} \times \frac{12}{37} + \frac{8}{38} \times \frac{7}{37} + \frac{1}{38} \times \frac{0}{37}$$

$P(2 \text{ students from same class})$

$$= \frac{4}{38} \times \frac{3}{37} + \frac{12}{38} \times \frac{11}{37} + \frac{13}{38} \times \frac{12}{37} + \frac{8}{38} \times \frac{7}{37} + \frac{1}{38} \times \frac{0}{37}$$

$$= 25.32\%$$

③ Same question as above but a different class

This class has: 400 Freshmen, 1200 sophomores, 1300 juniors, 800 seniors, 100 non-matriculated.

$$P(2 \text{ students from same class}) = P(2F) + P(2S) + P(2J) + P(2Sr) + P(2NM)$$

$$= \frac{400}{3800} \times \frac{399}{3799} + \frac{1200}{3800} \times \frac{1199}{3799} + \frac{1300}{3800} \times \frac{1299}{3799} + \frac{800}{3800} \times \frac{799}{3799} + \frac{100}{3800} \times \frac{99}{3799}$$

$$P(F)P(F|F) = 27.27\%$$

For very large populations sampling with or without replacement is not an important factor.

For small populations, each sample is no longer independent of others unless we use replacement.