

Løsningsforslag ST2301 Øving 2

Kapittel 1

Exercise 6

Har et utvalg på 200 individer, fra en populasjon med forventet Hardy-Weinbergandeler for et locus med tre alleler, A_1 , A_2 og A_3 . Antall individer i utvalget av hver av de seks mulige genotypene er

Genotype	Antall
A_1A_1	76
A_1A_2	54
A_1A_3	33
A_2A_2	18
A_2A_3	16
A_3A_3	3

Finn genfrekvensene til de tre allelene, og forventet antall individer av hver genotype utfra disse. Uten å gjøre en statistisk test, er det noen åpenbare forskjeller mellom utvalget og forventningene?

Svar:

Genfrekvensene er

$$\begin{aligned}p_{A_1} &= \frac{2N_{A_1A_1} + N_{A_1A_2} + N_{A_1A_3}}{2N} = \frac{2 \cdot 76 + 54 + 33}{2 \cdot 200} = 0.5975 \\p_{A_2} &= \frac{2N_{A_2A_2} + N_{A_1A_2} + N_{A_2A_3}}{2N} = \frac{2 \cdot 18 + 54 + 16}{2 \cdot 200} = 0.2650 \\p_{A_3} &= \frac{2N_{A_3A_3} + N_{A_1A_3} + N_{A_2A_3}}{2N} = \frac{2 \cdot 3 + 33 + 16}{2 \cdot 200} = 0.1375\end{aligned}$$

Skriver forventet antall individer av hver genotype i en tabell, sammen med genotypfrekvens, og faktisk antall av hver genotype.

Genotype	Genotypefrekvens	Forventet antall	Faktisk antall
A_1A_1	$p_{A_1}^2$	71	72
A_1A_2	$2p_{A_1}p_{A_2}$	63	54
A_1A_3	$2p_{A_1}p_{A_3}$	33	33
A_2A_2	$p_{A_2}^2$	14	18
A_2A_3	$2p_{A_2}p_{A_3}$	15	16
A_3A_3	$p_{A_3}^2$	4	3

Uten å utføre en statistisk test er det vanskelig å avgjøre om utvalget kommer fra en populasjon med Hardy-Weinbergandeler.

Exercise 7

Ved et kjønnskoplet locus er frekvensen av aY -hemizygoter blant hanner lik 0,2, og frekvensen aa -homozygoter blant hunner lik 0,1. Anta tilfeldig parring. Hva var genfrekvensene i de to kjønnene forrige generasjon? Hva vil genotypefrekvensene bli neste generasjon?

Svar:

La p_m være frekvensen av allel A hos hanner, og p_f frekvensen av A hos hunner, forrige generasjon. Disse finner man ut fra genotypefrekvensen hos hanner og hunner i nåværende generasjon. For hanner har man (likning I-36 s. 15)

$$P_{aY} = 1 - p_f$$

$$p_f = 1 - 0.2 = 0.8$$

Genotypefrekvensene hos hunner (likning I-35 s.14) gir

$$P_{aa} = (1 - p_f)(1 - p_m)$$

$$p_m = 1 - \frac{0.1}{0.2} = 0.5$$

Genfrekvensene for nåværende generasjon er gitt ved likning I-37 s. 15.

$$p'_f = \frac{p_f}{2} + \frac{p_m}{2} = \frac{0.8}{2} + \frac{0.5}{2} = 0.65$$

$$p'_m = p_f = 0.8$$

Dette gir genotypefrekvensene

$$P'_{aY} = 1 - p'_f = 1 - 0.65 = 0.35$$

$$P'_{aa} = (1 - p'_f)(1 - p'_m) = (1 - 0.65)(1 - 0.8) = 0.07$$

Exercise 8

Ved et kjønnskoplet locus med to allel A og a , finner man at genotypefrekvensene hos hunner og hanner er

$$\begin{array}{ccc} AA & Aa & aa \\ \hline 0.95 & 0.04 & 0.01 \end{array}$$

$$\begin{array}{cc} A- & a- \\ \hline 0.94 & 0.06 \end{array}$$

Genfrekvensene hos foreldrene til denne generasjonen er ikke nødvendigvis like i de to kjønnene.

1. Ser det ut som populasjonen er resultat av minst en generasjon med tilfeldig parring?
2. Hvis populasjonen reproducerer med tilfeldig parring en generasjon, hvilke genfrekvenser forventer man å se?

Svar:

1. Det finnes flere ulike måter å sjekke om populasjonen kan være resultat av tilfeldig parring. Man kan f.eks se om man finner samme p_m og p_f for ulike genotyper når man antar Hardy-Weinbergandeler.

$$P_{AY} = p_f = 0.94$$

$$P_{AA} = p_m p_f$$

$$p_m = 1.01$$

$$P_{aY} = 1 - p_f = 0.06$$

$$P_{aa} = (1 - p_m)(1 - p_f)$$

$$p_m = 0.83 \neq 1.01$$

Dvs. populasjonen virker ikke å være resultat av tilfeldig parring.

2. Genfrekvensene er gitt ved (likning I-8 s.4)

$$p_f = P_{AA} + \frac{1}{2}P_{Aa} = 0.97$$

$$p_m = P_{AY} = 0.94$$

Når man antar tilfeldig parring er genotypefrekvensene neste generasjon

gitt ved

$$\begin{aligned}P'_{AY} &= p_f = 0.97 \\ P'_{aY} &= 1 - p_f = 0.03\end{aligned}$$

$$\begin{aligned}P_{AA} &= p_f p_m = 0.9118 \\ P_{Aa} &= p_f(1 - p_m) + p_m(1 - p_f) = 0.0864 \\ P_{aa} &= (1 - p_m)(1 - p_f) = 0.0018\end{aligned}$$

Exercise 10

Ser på to loci, hver med to allel (A og a , B og b), i en populasjon med tilfeldig parring. Har $p_A = p_B = 0.5$, og $D_{AB} = 0.2$. Halvparten av individene er hunner og halvparten hanner. Rekombinasjonsraten mellom lociene er 0.3 for hunner og 0.1 for hanner.

1. Finn D'_{AB} i avkomsgenerasjonen som funksjon av D_{AB} nåværende generasjon.
2. Hva er frekvensen av genotype $AABB$ i avkomsgenerasjonen?

Svar:

1. Bruker loven om total sannsynlighet for å finne sannsynligheten for rekombinasjon i en tilfeldig gamet (uavhengig av kjønn). La

$$\begin{aligned}r &= P(\text{Rekombinasjon}|m)P(m) + P(\text{Rekombinasjon}|f)P(f) \\ &= r_m \cdot 0.5 + r_f \cdot 0.5 \\ &= 0.1 \cdot 0.5 + 0.3 \cdot 0.5 \\ &= 0.2\end{aligned}$$

Likning I-44 s.18 gir sammenhengen

$$\begin{aligned}D'_{AB} &= (1 - r)D_{AB} \\ &= 0.8D_{AB}\end{aligned}$$

2. For nåværende generasjon er

$$P_{AB} = p_A p_B + D_{AB} = 0.45$$

For å finne frekvensen av genotype $AABB$ neste generasjon trenger man gametfrekvensen P'_{AB} . Denne vil imidlertid ikke være den samme for begge kjønn, fordi rekombinasjonsraten er forskjellig. Bruker likning I-42 s.18 for

å finne gametfrekvensen for hvert kjønn.

$$\begin{aligned}P'_{AB}(m) &= (1 - r_m)P_{AB} + r_m p_A p_B \\ &= (1 - 0.1) \cdot 0.45 + 0.1 \cdot 0.5 \cdot 0.5 \\ &= 0.43\end{aligned}$$

$$\begin{aligned}P'_{AB}(f) &= (1 - r_f)P_{AB} + r_f p_A p_B \\ &= (1 - 0.3) \cdot 0.45 + 0.3 \cdot 0.5 \cdot 0.5 \\ &= 0.39\end{aligned}$$

Alle individer mottar en gamet fra far og en fra mor. Derfor er genotypfrekvensen for $AABB$ lik produktet av gametfrekvensene til AB hos hanner og hunner.

$$\begin{aligned}P'_{AABB} &= P'_{AB}(m)P'_{AB}(f) \\ &= 0.43 \cdot 0.39 \\ &= 0.1677\end{aligned}$$

Complement 1

Har et locus med n allel p_1, p_2, \dots, p_n . Etter tilfeldig parring,

1. Hvilken andel av kopiene av A_i inngår i heterozygoter?
2. Hva er totalandelen av kopier som inngår i heterozygoter?

Svar:

1. Det er N individer, og totalt $2Np_i$ kopier av allel A_i . Antall kopier av A_i i homozygoter er

$$2Np_i^2$$

Andelen kopier av A_i i heterozygoter er lik en minus andelen i homozygoter, dvs

$$1 - \frac{2Np_i^2}{2Np_i} = 1 - p_i$$

Alternativt kan man finne antall kopier av A_i blant heterozygoter,

$$\begin{aligned}2N \sum_{i \neq j}^N p_i p_j &= 2Np_i + \sum_{j=1}^{n-1} p_j \\ &= 2Np_i(1 - p_i)\end{aligned}$$

Andel kopier A_i som er i heterozygoter blir da

$$\frac{2Np_i(1-p_i)}{2Np_i} = 1 - p_i$$

2. Totalt antall allel er $2N$, og totalt antall allel i homozygoter er

$$\left(\sum_{i=1}^N p_i^2 \right) 2N$$

Andelen allel i homozygoter er

$$\frac{\left(\sum_{i=1}^N p_i^2 \right) 2N}{2N} = \sum_{i=1}^n p_i$$

Andelen allel i heterozygoter er derfor $1 - \sum_{i=1}^n p_i$.

Complement 3

Har et locus med to alleler A og a , som er koplet med et kjønnsbestemmende locus (kjønn 1 og 2) i en haploid organisme med tilfeldig parring. Rekombinasjonsraten mellom locuset og kjønnslocuset er r . Dersom initialfrekvensene av A er p_1 i det ene kjønnet og p_2 i det andre,

1. Hva er frekvensene av A neste generasjon?
2. Hva er frekvensene til A om t generasjoner?
3. Hva er de ultimate verdiene av p_1 og p_2 (hint: prøv å endre variabler og se på gjennomsnitt og differens mellom p_1 og p_2)?

Svar:

1. Ser på kjønn 1. En andel $(1-r)$ av gametene går gjennom formeringen en uten rekombinasjon, av disse har en andel p_1 allel A . En andel r av gametene gjennomgår en rekombinasjon. For at disse skal ende opp med allel A , må de parres med individer fra motsatt kjønn som har dette allelet (andel p_2). Neste generasjon er derfor

$$p'_1 = rp_2 + (1-r)p_1$$

Tilsvarende argument for kjønn 2 gir

$$p'_2 = rp_1 + (1-r)p_2$$

2. Tar hintet og definerer nye variabler:

$$\bar{p} = \frac{1}{2}(p_1 + p_2)$$
$$p^* = p_1 - p_2$$

$$p_1 = \bar{p} + \frac{1}{2}p^*$$

$$p_2 = \bar{p} - \frac{1}{2}p^*$$

Ser på hvordan gjennomsnittet endres i løpet av en generasjon.

$$\begin{aligned}\bar{p}' &= \frac{1}{2}(p_1' + p_2') \\ &= \frac{1}{2}(rp_2 + (1-r)p_1) + \frac{1}{2}(rp_1 + (1-r)p_2) \\ &= \frac{1}{2}(rp_2 + p_1 - rp_1 + rp_1 + p_2 - rp_2) \\ &= \frac{1}{2}(p_1 + p_2)\end{aligned}$$

Gjennomsnittet endrer seg ikke over tid, dvs $\bar{p}(t) = \bar{p}$. Ser deretter på hvordan differansen endrer seg i løpet av en generasjon.

$$\begin{aligned}p^* &= p_1' - p_2' \\ &= rp_2 + (1-r)p_1 - rp_1 - (1-r)p_2 \\ &= rp_2 + p_1 - rp_1 - rp_1 - p_2 + rp_2 \\ &= p_1 - p_2 - 2r(p_1 - p_2) \\ &= (1-2r)(p_1 - p_2)\end{aligned}$$

Det er to muligheter. Dersom $r = 0.5$ er differansen lik null. Er $r \neq 0.5$ avtar differansen med $(1-2r)$ per generasjon, dvs

$$p^*(t) = (1-2r)^t(p_1 - p_2)$$

Nå kan p_1 og p_2 uttrykkes ved t .

$$\begin{aligned} p_1(t) &= \bar{p}(t) + \frac{1}{2}p^*(t) \\ &= \frac{1}{2}(p_1 + p_2) + (1 - 2r)^t(p_1 - p_2) \end{aligned}$$

$$\begin{aligned} p_2(t) &= \bar{p}(t) - \frac{1}{2}p^*(t) \\ &= \frac{1}{2}(p_1 + p_2) + (1 - 2r)^t(p_2 - p_1) \end{aligned}$$

3. For å finne de ultimate verdiene av egenfrekvensene av A , må man la t gå mot uendelig. Det gir

$$\lim_{t \rightarrow \infty} (1 - 2r)^t = 0$$

og dermed

$$\begin{aligned} \lim_{t \rightarrow \infty} p_1(t) &= \frac{1}{2}(p_1 + p_2) \\ \lim_{t \rightarrow \infty} p_2(t) &= \frac{1}{2}(p_1 + p_2) \end{aligned}$$