FoS_HW_02_Group_II

Rehman Kerimov, Rene Plaß, Nils Beckmann

2025-10-30

Part I

Topic: Probability measures and distribution functions

Exercise 1

a) Define a new sequence $(B_n)_{n=1}^{\infty}$ as follows: $B_1 = A_1$, $B_n = A_n \setminus B_{n-1}$

Then all events B_n are pairwise disjoint and we get:

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{P}(A_n) = \lim_{N\to\infty} \mathbb{P}(A_N) = \lim_{n\to\infty} \mathbb{P}(A_n) \quad \blacksquare$$

b) If $(A_n)_{n=1}^{\infty}$ is an increasing sequence of events, then $(A_n^c)_{n=1}^{\infty}$ is increasing

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbb{P}\left(\left(\bigcap_{n=1}^{\infty} A_n\right)^{c^c}\right) = \mathbb{P}\left(\left(\bigcup_{n=1}^{\infty} A_n^c\right)^c\right)$$

$$= 1 - \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) \overset{a)}{=} 1 - \lim_{n\to\infty} \mathbb{P}(A_n^c) = \lim_{n\to\infty} \mathbb{P}(A_n) \quad \blacksquare$$

Exercise 2

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ be a random variable.

Definition: The cumulative distribution function (CDF) of X is defined by

$$F(x) = P(X \leq x) \quad \text{for all } x \in \mathbb{R}$$

a) Property 1: Monotonicity **Theorem:** F is monotonically increasing, i.e., for $x_1 < x_2$ we have $F(x_1) \leq F(x_2)$.

Proof:

$$A_1 = \{\omega \in \Omega : X(\omega) \leq x_1\}$$
$$A_2 = \{\omega \in \Omega : X(\omega) \leq x_2\}$$

Let $x_1 < x_2$. Then for the events:

$$A_1 \subseteq A_2$$

and thus

$$P(A_1) \leq P(A_2)$$

which is equal to

$$F(x) \leq F(y)$$

$\square$

$$\bigcup_{n=1}^{\infty} A_n = \{X \leq n \text{ for at least one } n\} = \Omega$$

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right)$$

Also:

$$\lim_{n\to\infty} F(n) = P(\Omega) = 1 \quad \square$$

---

b) Property 2: Limits

Part a: Limit at $-\infty$ **Theorem:** $\lim_{x\to-\infty} F(x) = 0$

Proof:

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence with $x_n \to -\infty$ and $x_n \downarrow$ (e.g., $x_n = -n$).

Set $A_n = \{X \leq x_n\}$. Then:

- (A_n) is a decreasing sequence of events and we have continuity from above $A_n \downarrow$
- $\bigcap_{n=1}^{\infty} A_n = \emptyset$ (since $x_n \to -\infty$)

By continuity from above:

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(\emptyset) = 0$$

Therefore

$$\lim_{n\to\infty} F(x_n) = 0 \quad \square$$

Part b: Limit at $+\infty$ **Theorem:** $\lim_{x\to+\infty} F(x) = 1$

Proof:

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence with $x_n \to +\infty$ and $x_n \uparrow$ (e.g., $x_n = n$).

Set $A_n = \{X \leq x_n\}$. Then:

- (A_n) is an increasing sequence of events and we have continuity from below $A_n \uparrow$
- $\bigcup_{n=1}^{\infty} A_n = \Omega$ (since $x_n \to \infty$)

By continuity from below:

$$\lim_{n\to\infty} P(A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = P(\Omega) = 1$$

Therefore

$$\lim_{n\to\infty} F(x_n) = 1 \quad \square$$

---

c) Property 3: Right-Continuity **Theorem:** For all $x_0 \in \mathbb{R}$: $\lim_{x\to x_0^+} F(x) = F(x_0)$

$x \to x_0^+$ means x approaches x_0 from the right (from above)

Proof:

Let (x_n) be a sequence with $x_n \downarrow x_0$ (e.g., $x_n = x_0 + \frac{1}{n}$).

Set $A_n = \{X \leq x_n\}$ and $A = \{X \leq x_0\}$. Then:

- (A_n) is a decreasing sequence
- $\bigcap_{n=1}^{\infty} A_n = A$ (since $X(\omega) \leq x_n \forall n \Rightarrow X(\omega) \leq x_0$)

By continuity from above:

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P(A_n) = P(A) = F(x_0) \quad \square$$

---

d) F is left-limited

Statement to Prove A distribution function F is **left-limited** at each point $x \in \mathbb{R}$, i.e., the left-hand limit exists:

$$F(x^-) := \lim_{z\nearrow x} F(z) \text{ exists for all } x \in \mathbb{R}$$

Properties of Distribution Functions Recall that F has the following properties:

1. **Monotone increasing**: For all $x \leq y$, we have $F(x) \leq F(y)$
2. **Right-continuous**: For all $x \in \mathbb{R}$, $\lim_{z \searrow x} F(z) = F(x)$
3. **Limits**: $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$

Proof **Fix an arbitrary point $x \in \mathbb{R}$.**

Step 1: Define the set of left-hand values Consider the set:

$$M = \{F(z) : z < x\}$$

This is the set of all values that F takes at points to the left of x.

Step 2: M is bounded above **Claim:** M is bounded above.

Proof: For any $z < x$, the monotonicity of F gives us $F(z) \leq F(x)$.

Therefore, $F(x)$ is an upper bound for M.

Step 3: The supremum exists Since M is non-empty (F is defined on all of $\mathbb{R}$) and bounded above, by the **completeness axiom of** $\mathbb{R}$, the supremum exists:

$$L := \sup M = \sup\{F(z) : z < x\}$$

Step 4: L is the left-hand limit **Claim:** $L = \lim_{z \nearrow x} F(z) = F(x^-)$

Proof: We must show that for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all z with $x - \delta < z < x$:

$$|F(z) - L| < \varepsilon$$

Let $\varepsilon > 0$ be given.

Part A: By definition of supremum, $F(z) \leq L$ for all $z < x$.

Therefore: $0 \leq L - F(z)$ for all $z < x$.

Part B: By the characterization of supremum, for any $\varepsilon > 0$, there exists $z_0 < x$ such that:

$$L - \varepsilon < F(z_0) \leq L$$

(Otherwise, $L - \varepsilon$ would be a smaller upper bound, contradicting that L is the *least* upper bound.)

Part C: Choose $\delta := x - z_0 > 0$.

For any z with $x - \delta < z < x$, we have:

- $z_0 < z < x$ (by choice of δ)
- $F(z_0) \leq F(z) \leq L$ (by monotonicity and Part A)
- $L - \varepsilon < F(z_0) \leq F(z) \leq L$ (combining with Part B)

Therefore:

$$L - \varepsilon < F(z) \leq L$$

which gives us:

$$0 \leq L - F(z) < \varepsilon$$

Hence:

$$|F(z) - L| < \varepsilon$$

This proves that $\lim_{z \nearrow x} F(z) = L$.

Conclusion

We have shown that the left-hand limit $F(x^-) = \lim_{z \nearrow x} F(z)$ exists and equals:

$$F(x^-) = \sup\{F(z) : z < x\} \quad \square$$

d) Counterexample: F is not necessarily left-continuous **Given:** F is left-limited, meaning the left limit $F(x^-) = lim_{z \nearrow x}\ F(z)$ exists at every point $x \in \mathbb{R}$.

To show: F can satisfy $F(x) \neq F(x^-)$, i.e., F is not necessarily left-continuous.

Counterexample:

Consider the CDF:

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

This is the CDF of a point mass at x = 0.

Evaluation at x = 0:

Left limit:

$$F(0^-) = \lim_{z \nearrow 0} F(z) = \lim_{z \nearrow 0} 0 = 0$$

(since F(z) = 0 for all z < 0)

Function value:

$$F(0) = 1$$

Conclusion:

$$F(0^-) = 0 \neq 1 = F(0)$$

Therefore, F is left-limited (the left limit exists) but **not left-continuous** at x = 0.

This demonstrates that a CDF can have $F(x) \neq F(x^-) \quad \square$.

(e) For $x \leq y$:

i)

$$\{X \leq y\} = \{X \leq x\} \dot\cup \{x < X \leq y\} \quad \Longrightarrow \quad P(x < X \leq y) = F(y) - F(x)$$

ii)

$$\{X \leq y\} = \{X < x\} \dot\cup \{x \leq X \leq y\} \quad \Longrightarrow \quad P(x \leq X \leq y) = F(y) - P(X < x) = F(y) - F(x^-)$$

iii)

$$\{X \leq x\} = \{X < x\} \dot\cup \{X = x\} \quad \Longrightarrow \quad P(X = x) = F(x) - F(x^-)$$

Exercise 3

3 a) $\mathbb{Q}(\Omega) = \mathbb{P}(\Omega | B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} \underset{B \subseteq \Omega}{=} \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$ ✓

b) Because $(A_n)_{n=1}^{\infty}$ mutually disjoint

$(A_n | B)_{n=1}^{\infty}$ have to be disjoint

~~$C_n = A_n | B$ $((C_n)_{n=1}^{\infty}$ mutually disjoint and countable~~

$\Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n | B) = \frac{\mathbb{P}(\bigcup_{n=1}^{\infty} A_n \cap B)}{\mathbb{P}(B)} = \sum_{n=1}^{\infty} \frac{\mathbb{P}(A_n \cap B)}{\mathbb{P}(B)}$

$= \sum_{n=1}^{\infty} \mathbb{P}(A_n | B)$ $\square$

c) $A \cup A^c \overset{(*)}{=} \Omega$

$\mathbb{P}(A|B) + \mathbb{P}(A^c|B) \overset{(b)}{=} \mathbb{P}(A|B \cup A^c|B)$

$\overset{(*)}{=} \mathbb{P}(\Omega | B) = \mathbb{Q}(\Omega) \overset{(a)}{=} 1$

Alternativ

$\mathbb{P}(A|B) + \mathbb{P}(A^c|B)$ ~~$= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A \cap B) \cup (A^c \cap B)}$~~

$= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} + \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((A \cap B) \cup (A^c \cap B))}{\mathbb{P}(B)} = \frac{\mathbb{P}((A \cup A^c) \cap B)}{\mathbb{P}(B)}$

$= \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$

Exercise 4

4) (i) and (ii) Are the same so 7

a show it ones

$$\overbrace{(A^c \cap B) \cup (A \cap B)}^{\text{Disjoint}} = B$$

$\Rightarrow$

$$P(B) = P(A^c \cap B) + P(A \cap B)$$

$$\Leftrightarrow P(B) = P(A^c \cap B) + P(A) \cdot P(B)$$

$$\Leftrightarrow P(A^c \cap B) = P(B) - P(A) \cdot P(B) = \overbrace{[1 - P(A)]}^{P(A^c)} \cdot P(B)$$

~~A ist~~ Case 1 $A = A_1$ Case 2 $A = A_2$

iii)

A_1 independent of $A_2 \overset{i)}{\Rightarrow} A_1^c$ ind. of A_2

$\overset{ii)}{\Rightarrow}$ A_1^c ind of A_2^c $\square$

$A_1^c = A$

4 c) Assumtion $1 - x \overset{*}{\leq} e^{-x}$

Esy to see

slop of -1

slop

is ~~stricte~~ bigger than -1 $x > 0$

$$\Rightarrow \mathbb{P}\left(\bigcap_{n=1}^{N} A_n^c\right) \overset{4b)}{=} \prod_{n=1}^{N} \mathbb{P}(A_n^c) = \prod_{n=1}^{N} 1 - \mathbb{P}(A_n)$$

$e^a \cdot e^b = e^{a+b}$

$$\overset{*}{\leq} \prod_{n=1}^{N} e^{-\mathbb{P}(A_n)} = \exp\left\{-\sum_{n=1}^{N} \mathbb{P}(A_n)\right\} \quad \square$$

...

Part II

Exercise 5

a) false positive: $\mathbb{P}(+\mid N)$ → positive test, given that person is negativ.

$$1 = \mathbb{P}(+\mid N) + \mathbb{P}(-\mid N) \Leftrightarrow \mathbb{P}(+\mid N) = 1 - \mathbb{P}(-\mid N)$$

$$\Rightarrow \mathbb{P}(+\mid N) = 1 - 0{,}9968 = 0{,}0032$$

b) $$\mathbb{P}(+) = \mathbb{P}(+\mid C)\cdot\mathbb{P}(C) + \mathbb{P}(+\mid N)\,\mathbb{P}(N)$$

$$= 0.9652\cdot\frac{7}{100\,000} + 0.0032\cdot\frac{99\,993}{100\,000} \approx 0.00327$$

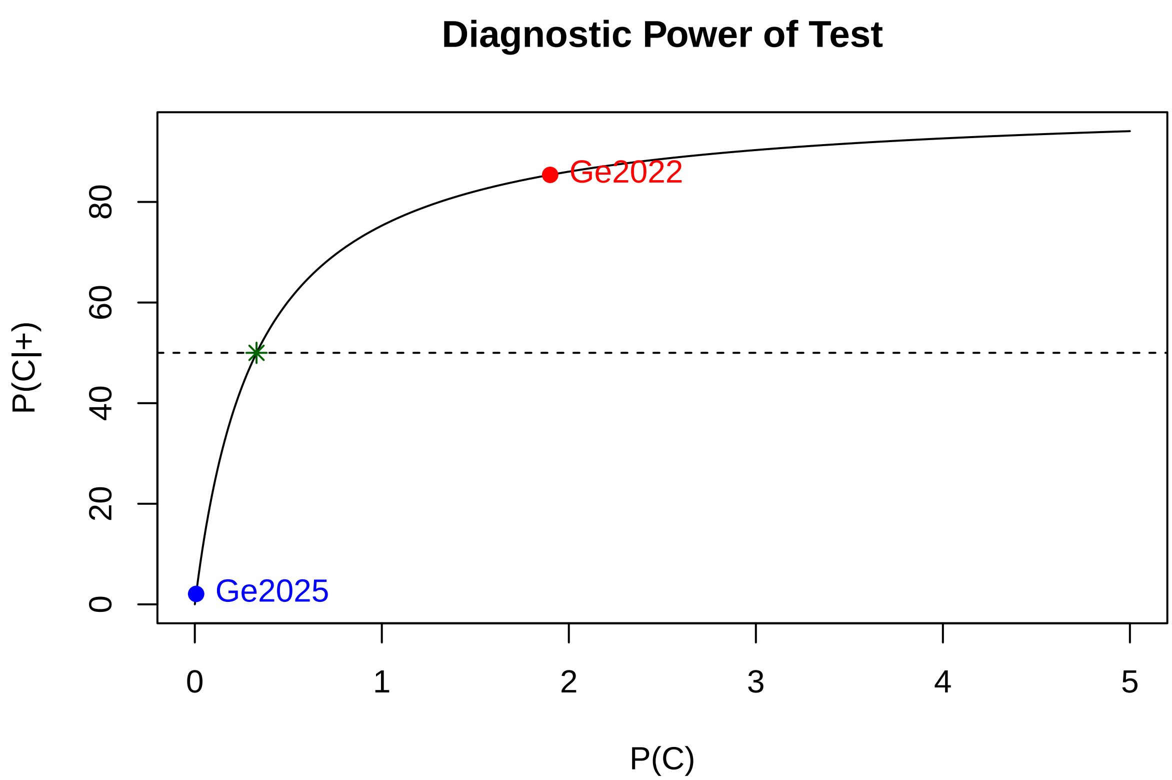
c) $$\mathbb{P}(C\mid +) = \frac{\mathbb{P}(+\mid C)\cdot\mathbb{P}(C)}{\mathbb{P}(+)} \approx \frac{0.9652\cdot\frac{7}{100000}}{0.00327} \approx 0{,}0207$$

d) $$\mathbb{P}(C\mid -) = \frac{\mathbb{P}(-\mid C)\cdot\mathbb{P}(C)}{\mathbb{P}(-)} = \frac{(1-\mathbb{P}(+\mid C))\cdot\mathbb{P}(C)}{(1-\mathbb{P}(+))} =$$

$$\approx \frac{0{,}0348\cdot\frac{7}{100\,000}}{0{,}99\,673} \approx 0{,}0000024$$

```r
# Given parameters
sens <- 0.9652
spec <- 0.9968
prob_false_pos <- 1 - spec

# Incidence rates (0% to 5%)
prob_C <- seq(0, 0.05, by = 0.0001)

# Compute P(C|+)
prob_C_given_pos <- (sens * prob_C) / (sens * prob_C + prob_false_pos * (1 - prob_C))

# Plot
```

```r
plot(prob_C * 100, prob_C_given_pos * 100, type = "l",
     xlab = "P(C)",
     ylab = "P(C|+)",
     main = "Diagnostic Power of Test")

points(0.00007 * 100, 0.0207 * 100, col = "blue", pch = 19)
text(0.00007 * 100, 0.0207 * 100, "Ge2025", pos = 4, col = "blue")

points(0.019 * 100, 0.8538 * 100, col = "red", pch = 19)
text(0.019 * 100, 0.8538 * 100, "Ge2022", pos = 4, col = "red")

abline(h = 50, lty = 2)
prob_C_given_pos_50 <- 0.0033
points(prob_C_given_pos_50 * 100, 50, pch = 8, col = "darkgreen", cex = 1)
```



e), f)

```r
set.seed(420)
n <- 1000000
prob_C <- 0.00007

# Simulate infection status
C <- rbinom(n, 1, prob_C)
# Simulate test results
```

```
test_pos <- ifelse(C == 1,
                   rbinom(n, 1, sens),
                   rbinom(n, 1, 1 - spec))
# Estimate P(C|+)
P_C_given_pos_sim <- mean(C[test_pos == 1])
P_C_given_pos_sim
```

g)

```
## [1] 0.01850139
```

Topic: The Hardy-Weinberg law in population genetics

Part III

Aufgabe 6

a) Searched is; $P(Aa \cdot Aa \cap aa)$ Mating is independent:

$$P(Aa \cdot Aa) = 2q \cdot 2q = 4q^2$$

Offspring being aa:

$$P(aa|Aa \cdot Aa) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$P(A \cap B) = P(A) \cdot P(B|A)$:

$$P(Aa \cdot Aa \cap aa) = P(Aa \cdot Aa) \cdot P(aa|Aa \cdot Aa) = 4q^2 \cdot \frac{1}{4} = q^2$$

b) The question arises as to how much the individual genotypes contribute to A.

- AA: (guaranteed) A Allele $\rightarrow$ p
- Aa: half A Allele $\rightarrow$ q
- aa: zero A Allele $\rightarrow$ 0

Thus

- Total frequency of A: $p + q$
- analog for a: $r + q$

Offspring having AA in the first generation p_1

$$(p + q) \cdot (p + q) = (p + q)^2$$

Offspring having Aa in the first generation $2q_1$

$$(p + q) \cdot (r + q) + (r + q) \cdot (p + q) = 2(r + q) \cdot (p + q)$$

Offspring having aa in the first generation r_1

$$(r + q) \cdot (r + q) = (r + q)^2$$

(c) Show: $p_2 = p_1$, $2q_2 = 2q_1$, $r_2 = r_1$

Frequency of A in generation 1:

$$p_1 + q_1 = (p+q)^2 + (p+q)(r+q) = (p+q)\underbrace{((p+q)+(r+q))}_{=1} = (p+q)\cdot 1 = p+q$$

Frequency of a in generation 1:

$$r_1 + q_1 = (r+q)^2 + (p+q)(r+q) = (r+q)[(r+q)+(p+q)] = (r+q)\cdot 1 = r+q$$

Key observation

The Allele frequencies in generation 1 are the **same** as in generation 0.

Computing generation 2

Since Allele frequencies are unchanged, applying the same logic as in part (b):

$$p_2 = (p+q)^2 = p_1$$
$$2q_2 = 2(p+q)(r+q) = 2q_1$$
$$r_2 = (r+q)^2 = r_1$$

(d) Show: $p_n = p_1$, $2q_n = 2q_1$, $r_n = r_1$ for all $n \geq 1$

Proof by induction:

Base case: $n = 1$

Trivially true: $p_1 = p_1$, $2q_1 = 2q_1$, $r_1 = r_1$✓

Inductive hypothesis: Assume for some $k \geq 1$:

$$p_k = p_1, \quad 2q_k = 2q_1, \quad r_k = r_1$$

Inductive step: Show $p_{k+1} = p_1$, $2q_{k+1} = 2q_1$, $r_{k+1} = r_1$.

From part (c), we know that if generation k has frequencies $p_k, 2q_k, r_k$, then its allele frequencies are:

$$p_k + q_k = p_1 + q_1 = p + q$$

$$r_k + q_k = r_1 + q_1 = r + q$$

Applying random mating to generation k:

$$p_{k+1} = (p_k + q_k)^2 = (p+q)^2 = p_1$$

$$2q_{k+1} = 2(p_k+q_k)(r_k+q_k) = 2(p+q)(r+q) = 2q_1$$
$$r_{k+1} = (r_k+q_k)^2 = (r+q)^2 = r_1$$

By induction, the result holds for all $n \geq 1$. □

Hardy-Weinberg Law

Conclusion: After one generation of random mating, genotype frequencies reach equilibrium and remain constant forever:

$$\boxed{p_n = (p+q)^2, \quad 2q_n = 2(p+q)(r+q), \quad r_n = (r+q)^2 \quad \forall n \geq 1}$$

Implications:

(i) No evolutionary change occurs through reproduction alone.

(ii) Changes in allele/genotype frequencies require additional forces:

- Natural selection
- Genetic drift
- Mutations
- Migration