Machine Learning Exam - Bilet 1 Task Solutions

# Task 1: Variance of a Sum

In this task, we prove that the variance of the sum of two random variables is given by:  
var(X + Y) = var(X) + var(Y) + 2cov(X, Y), where cov(X, Y) is the covariance between X and Y.

Proof:

1. Variance definition: var(Z) = E[(Z - E[Z])^2]  
2. var(X + Y) = E[(X + Y - E[X + Y])^2] = E[(X - E[X] + Y - E[Y])^2]  
3. Expanding the square: var(X + Y) = var(X) + var(Y) + 2cov(X, Y)  
4. Final result: var(X + Y) = var(X) + var(Y) + 2cov(X, Y)

# Task 2: Conditional Independence

We are asked to calculate P(H|e1, e2) given three sets of numbers and explore conditional independence.   
Using Bayes' Rule, the required set of numbers simplifies when E1 ⊥ E2 | H.

Sufficient Sets for Calculation:

- Using Bayes Rule: P(H|e1, e2) = P(e1, e2|H)P(H) / P(e1, e2)  
- Conditional independence simplifies this to: P(H|e1, e2) = P(e1|H)P(e2|H)P(H).  
- Final result: The set {P(e1|H), P(e2|H), P(H)} is sufficient for calculating P(H|e1, e2).

# Task 3: Conditional Independence if Joint Factorizes

We are asked to prove the alternative definition of conditional independence X ⊥ Y | Z if P(x, y|z) = g(x, z)h(y, z).

Proof:

1. Conditional independence definition: P(x, y | z) = P(x | z) P(y | z)  
2. Assume the joint factorizes as P(x, y | z) = g(x, z) h(y, z)  
3. Setting g(x, z) = P(x | z) and h(y, z) = P(y | z), we satisfy the definition of conditional independence.  
4. Final result: We have proven the alternative definition.