***למידה מחיזוקים – מטלה מספר 4***

***מגישים:*** *אופק ברנסקי, דן מלכה, דן ברוך*

**שאלה 1:**

1. We will bound from above the sum of probabilities of playing arm 2 at times when event (i) occurs, using Hoeffding’s inequality:
2. We will bound from above the sum of probabilities of always playing arm 2 when event (ii) occurs, by using the relevant algorithm definition & Hoefding’s bound:

*\* The last two equations are based on our calculations from Question a,   
and Hoefding’s bound.*

1. We will calculate the regret using only :

**שאלה 2:**

1. פונקציית ההסתברות המצטברת עבור התפלגות מעריכית היא עבור .

לכן, ההסתברות לדגום היא

.

1. נחשב את :
2. עדכון REINFORCE בהינתן יהיה: