MA323 Lab08 Report

Pavan Kumar A
210123043

Q1)

Strata:
S=0, S = 1, S=2, S=3, S=4, S=5, S>=6

For stratum S = 0, we know for sure that the total rainfall(X) is X<5
So, we take only 2 samples from that stratum.
We decide the number of samples for the other strata in the proportion to P(S = s) / (1 – P(S=0)).

We get the following number of sample count for the strata for different values of n.

For N = 100

S	0	1	2	3	4	5	>=6
n	2	16	24	23	17	9	9

For N = 10000

S	0	1	2	3	4	5	>=6
n	2	1668	2448	2366	1715	995	786

Results:
Monte Carlo Simulation:

N	Estimated Probability	99% Confidence Interval
100	0.38	(0.254973, 0.505027)
10000	0.3732	(0.360742, 0.385658)

Stratification Simulation:

N	Estimated Probability	99% Confidence Interval
100	0.334378	(0.212857, 0.455898)
10000	0.373266	(0.360808, 0.385725)

Q2)

$$f(x_1, \ldots, x_{38}) = \begin{cases} \frac{\Gamma(\sum_{j=1}^{38} \alpha_j)}{\prod_{j=1}^{38} \Gamma(\alpha_j)} \prod_{j=1}^{38} x_j^{\alpha_j - 1} & \text{if } x_j > 0 \text{ and } \sum_{i=1}^{38} x_i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

j	1	2	3	4	5	6	7	8	9	10	11	12	13
α_j	2082	1999	2008	2047	2199	2153	1999	2136	2053	2121	1974	2110	2110
j	14	15	16	17	18	19	20	21	22	23	24	25	26
α_j	2168	2035	2019	2044	2191	2284	1912	2196	2099	2041	2192	2188	1984
j	27	28	29	30	31	32	33	34	35	36	37	38	
α_j	2158	2019	2032	2051	2192	2133	2142	2113	2150	2221	2046	2127	

To estimate μ we repeatedly sample $Y \in [0, \infty)^{38}$ and averaging $f(Y)$. Here we condition on Y_{19}.

Given that $Y_{19} = y_{19}$, the probability that Y_{19} is largest is

$h(y_{19}) = \prod_{j=1, j\neq 19}^{38} G_{\alpha j} (y_{19})$

where $G_\alpha (x) = {}_0\int^x e^{-y}\, y^{\alpha-1}\, dy/\Gamma(\alpha)$

The required probability estimate is 0.6287263668298718 in this case

Q3)

The difference estimator is given by:

$$\hat{\mu}_{\text{diff}} = \frac{1}{n}\sum_{i=1}^{n}\big(f(\boldsymbol{X}_i) - h(\boldsymbol{X}_i)\big) + \theta = \hat{\mu} - \hat{\theta} + \theta.$$

where $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f\left(X_i\right)$ and $\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$. $\theta = E(h(\boldsymbol{X}))$

β optimal given by:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}\left(f(\boldsymbol{X}_i) - \widehat{\mu}\right)\left(h(\boldsymbol{X}_i) - \widehat{\theta}\right)}{\sum_{i=1}^{n}\left(h(\boldsymbol{X}_i) - \widehat{\theta}\right)^2},$$

which is estimated as 4.728878352837489e-07
For 10000 samples, Estimated mean = 12.113945174400321 with 99% confidence interval - (8.140334157192198, 16.087556191608442)
With Parameters: (0, 5), (0, 5), (0, 5), (0, 5), (0, 5)