STATS 551 - HW 5 - Q1

```r
# Data
y <- c(16/74, 9/99, 10/58, 13/70, 19/121, 20/77, 18/104, 17/129, 35/308, 55/119)
z <- c(12/25, 1/19, 2/16, 4/48, 9/217, 7/74, 9/38, 8/162)

# Total counts
n_y <- c(74, 99, 58, 70, 121, 77, 104, 129, 308, 119)
n_z <- c(25, 19, 16, 48, 217, 74, 38, 162)

# Prior parameters
alpha_y <- 1; beta_y <- 1
alpha_z <- 1; beta_z <- 1

# Posterior parameters
alpha_y_post <- alpha_y + sum(y * n_y)
beta_y_post <- beta_y + sum((1 - y) * n_y)
alpha_z_post <- alpha_z + sum(z * n_z)
beta_z_post <- beta_z + sum((1 - z) * n_z)

# Simulate from the posterior
set.seed(123)
theta_y_samples <- rbeta(1000, alpha_y_post, beta_y_post)
theta_z_samples <- rbeta(1000, alpha_z_post, beta_z_post)

# Compute the difference
diff_samples <- theta_y_samples - theta_z_samples

# Plot
hist(diff_samples, main = "Posterior Distribution of Difference",
     xlab = expression(theta[y] - theta[z]), breaks = 40, col = "skyblue")
```