Statistics with Recitation: TA Session

Danny Po-Hsien Kang (康柏賢)

November 4, 2025

Today's agenda

1. Random Number Generator
   - `rpois()`
   - `rexp()`
2. Data Wrangling: Combine Dataset
   - `rbind()`
   - `cbind()`

Generate Poisson Random Samples: rpois()

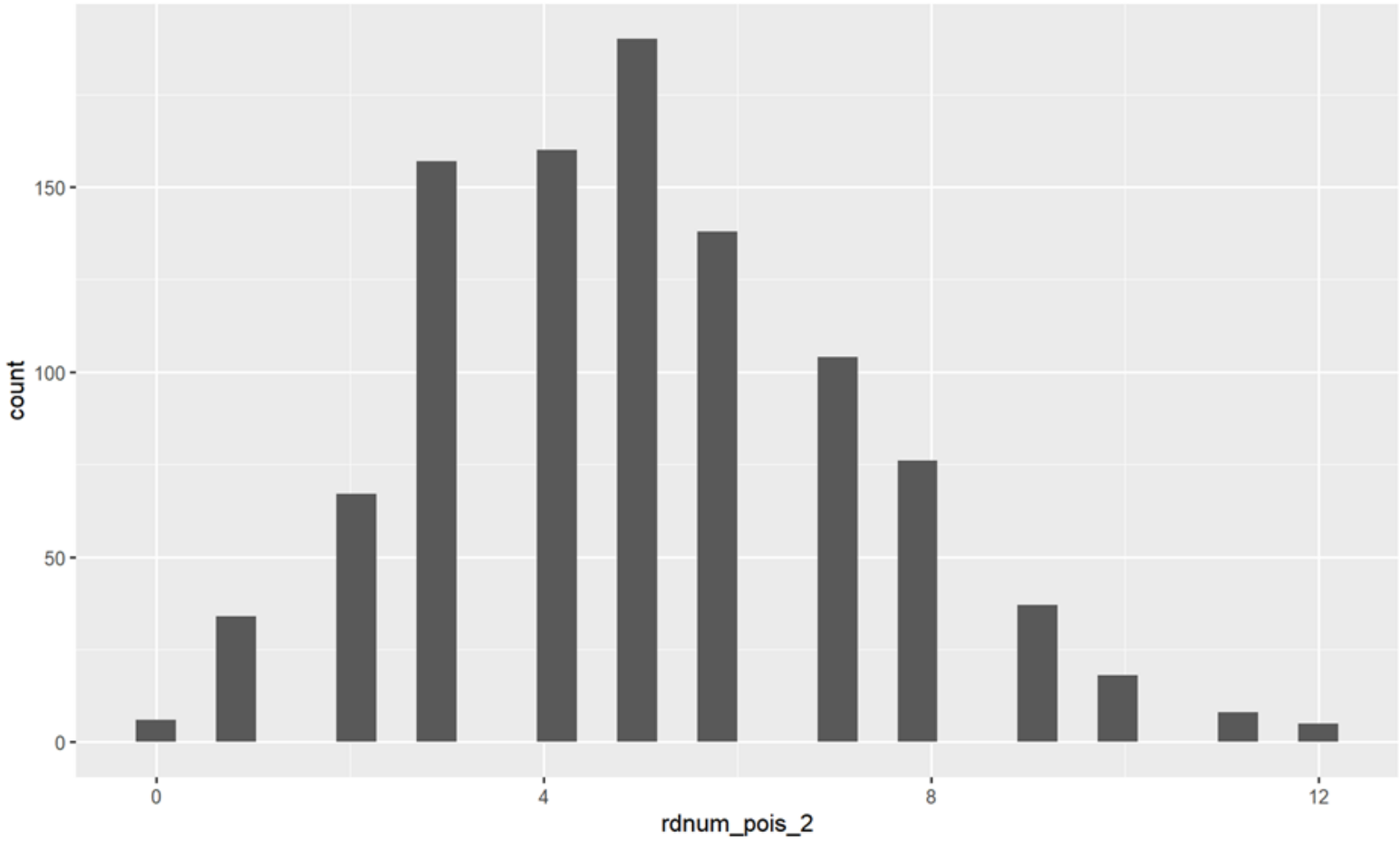

Figure: Random samples with size 1,000 drawn from Poisson($\lambda = 5$)

Generate Poisson Random Samples: rpois()

- **Syntax:**

```
num <- rpois(n = ..., lambda = ...)
```

- **Example:**

```
rdnum_pois <- rpois(20, lambda = 5)
```

- **Output:**

```
> print(rdnum_pois)
 [1] 6 5 5 5 7 5 8 3 6 5 4 9 6 1 5 7 4 4 6 2
```

Generate Exponential Random Samples: rexp()

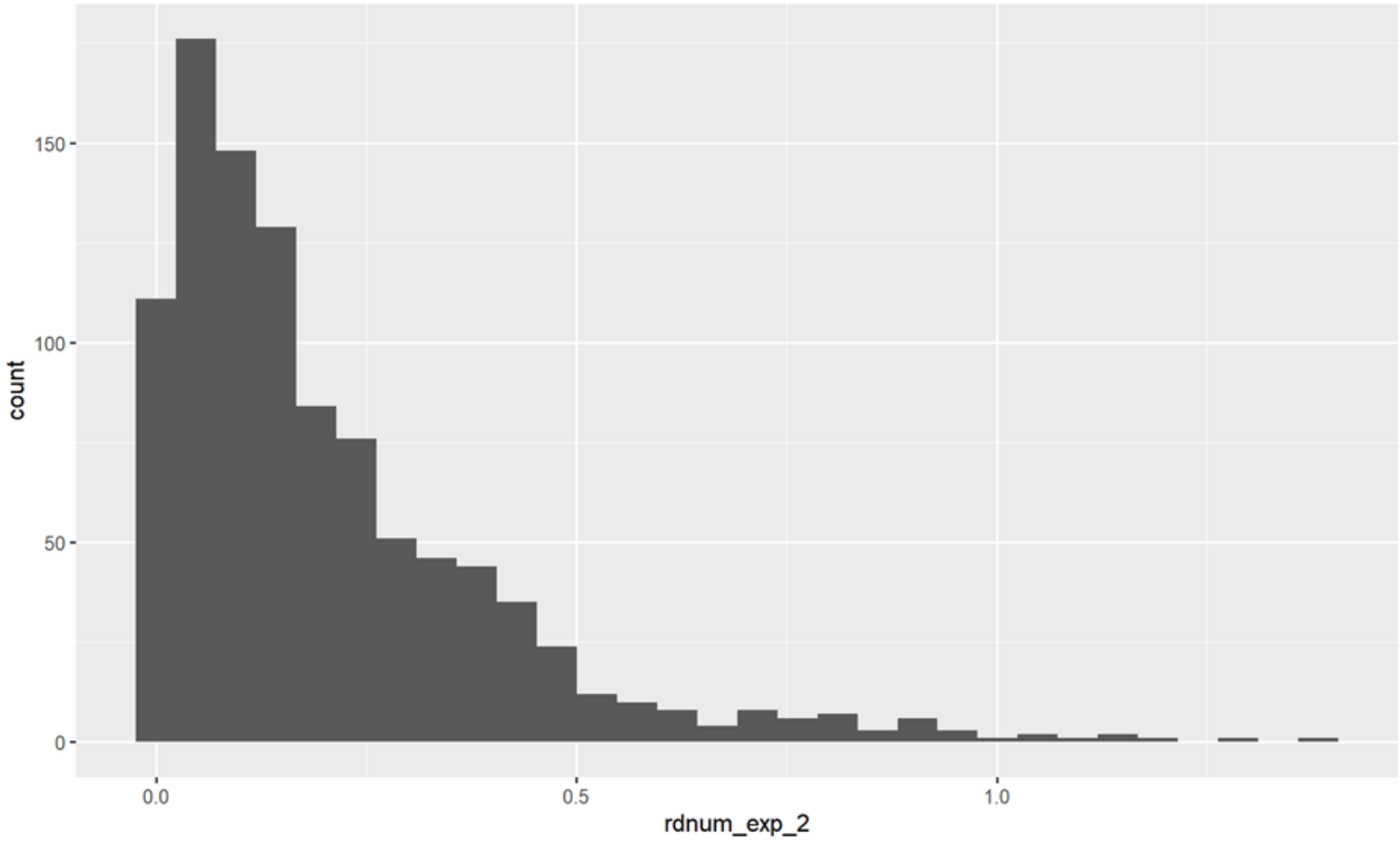

Figure: Random samples with size 1,000 drawn from Exp($\lambda = 5$)

Generate Exponential Random Samples: rexp()

- **Syntax:**

```
num <- rexp(n = ..., rate = ...)
```

- **Example:**

```
rdnum_exp <- rexp(20, rate = 5)
```

- **Output:**

```
> print(rdnum_exp)
 [1] 0.142843 0.273646 0.044956 0.259284 0.039657
 [6] 0.296780 0.413548 0.067782 0.022841 0.092401
[11] 0.449154 0.004849 0.161354 0.095610 0.189128
[16] 0.630804 0.006718 0.248812 0.022429 0.772203
```

Two Ways to Combine Your Dataset

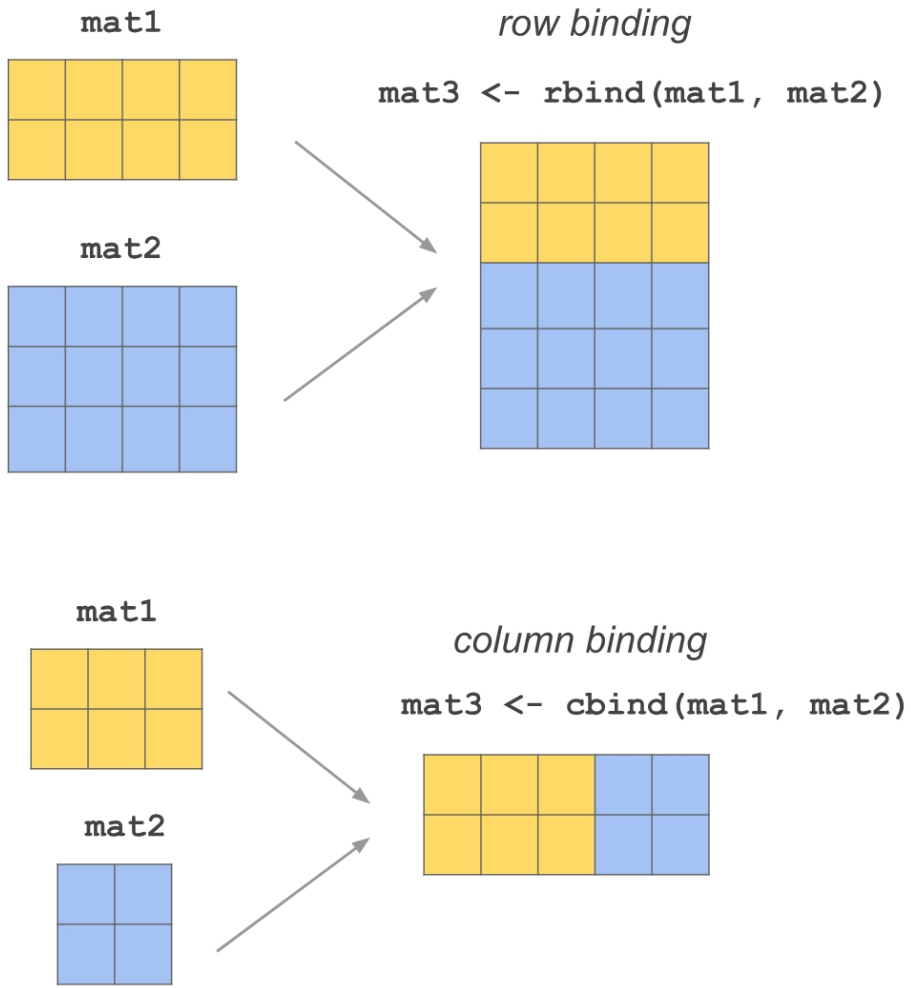

Figure: Combine dataset by rows or columns (Gaston Sanchez, 2024)

Combine Dataset by Rows: rbind()

- **Example:**

```
## first, let's draw two different sets of random sample
rdnum_pois_20 <- rpois(n = 20, lambda = 5)
rdnum_pois_10000 <- rpois(n = 10000, lambda = 5)

## using the two sets of random sample, build two DataFrames
df_20 <- data.frame(
    size = 20,
    values = rdnum_pois_20
)
df_10000 <- data.frame(
    size = 10000,
    values = rdnum_pois_10000
)

## combine the two DataFrames together, and draw the histogram
df <- rbind(df_20, df_10000)
```

Combine Dataset by Rows: rbind()

- **Example:**

```
ggplot(df, aes(x = values)) +
    geom_histogram() +
    facet_wrap(~size, scales = "free_y")
```

Combine Dataset by Rows: rbind()

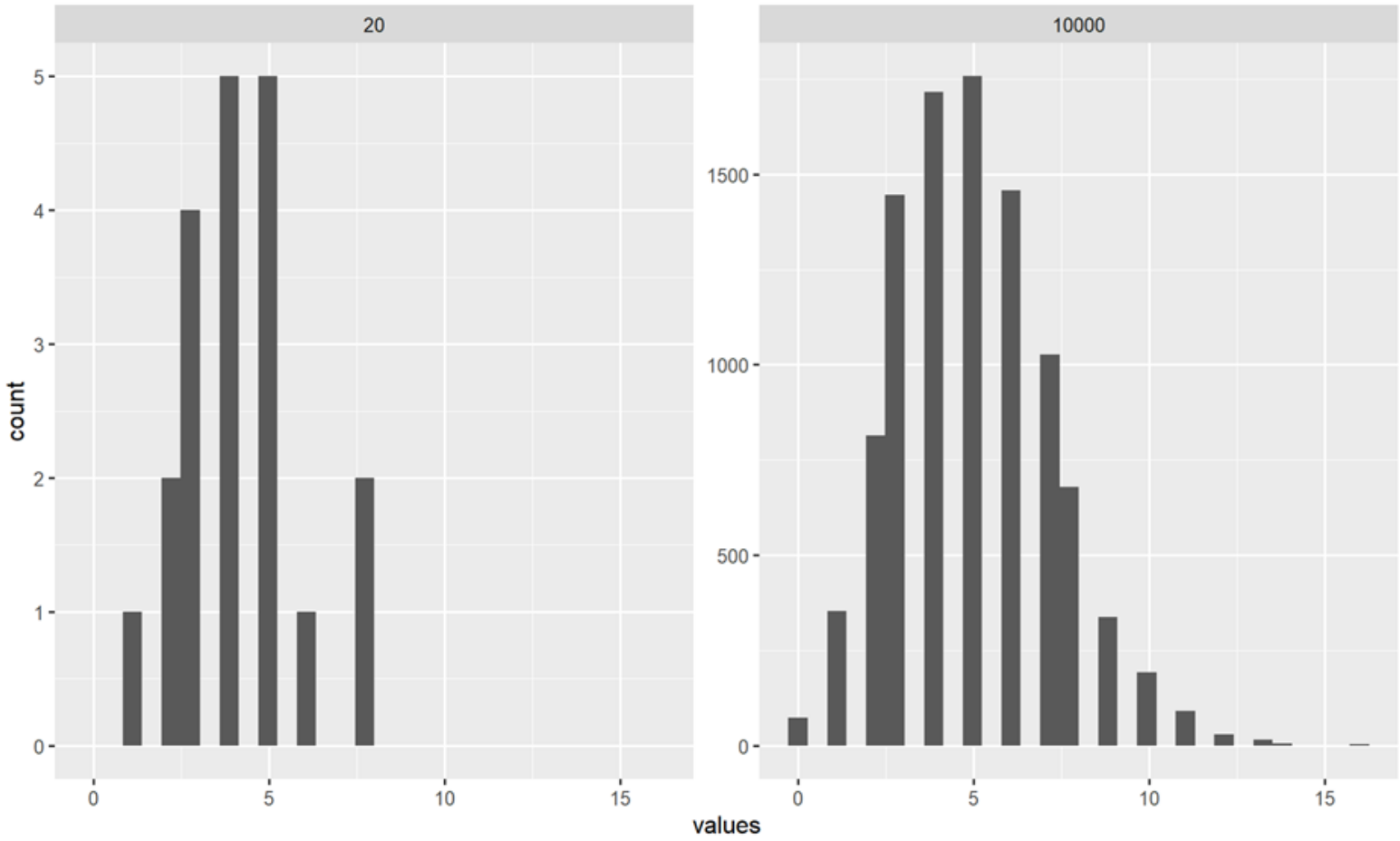

Figure: Histogram of two samples in different sample sizes

Combine Dataset by Columns: cbind()

- **Example:**

```
## first, let's draw two different sets of random sample
rdnum_pois_param5 <- rpois(n = 1000, lambda = 5)
rdnum_pois_param20 <- rpois(n = 1000, lambda = 20)

## using the two sets of random sample, build two DataFrames re
df_param5 <- data.frame(
    values_param5 = rdnum_pois_param5
)
df_param20 <- data.frame(
    values_param20 = rdnum_pois_param20
)
```

Combine Dataset by Columns: cbind()

- **Example:**

```
df <- cbind(df_param5, df_param20)

ggplot(df) +
    geom_histogram(mapping = aes(x = values_param5,
                                 fill = "lambda = 5"),
                   alpha = 0.7
    ) +
    geom_histogram(mapping = aes(x = values_param20,
                                 fill = "lambda = 20"),
                   alpha = 0.7
    )
```

Combine Dataset by Columns: cbind()

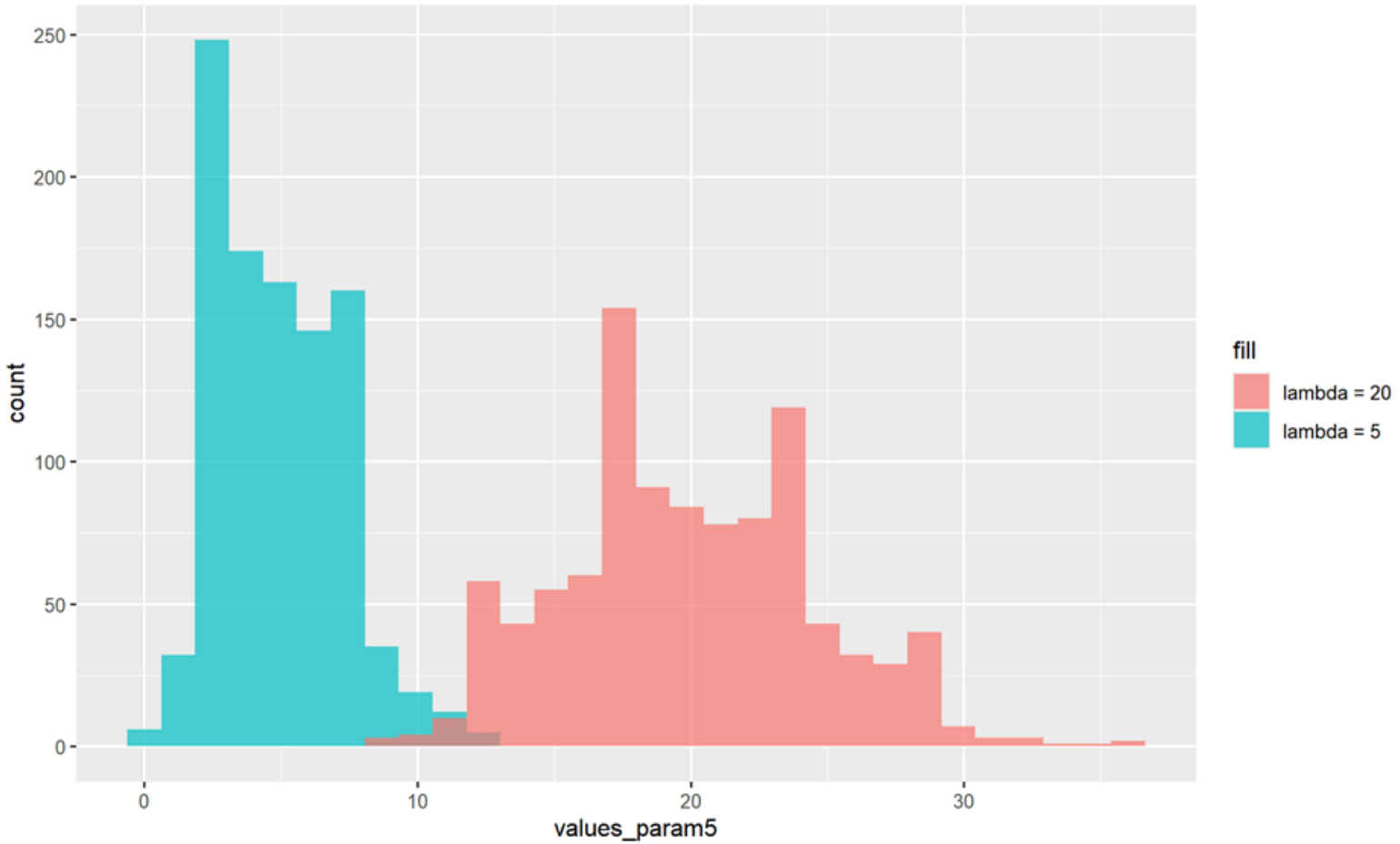

Figure: Histogram of two samples in different distribution parameters