R Notebook

General results:

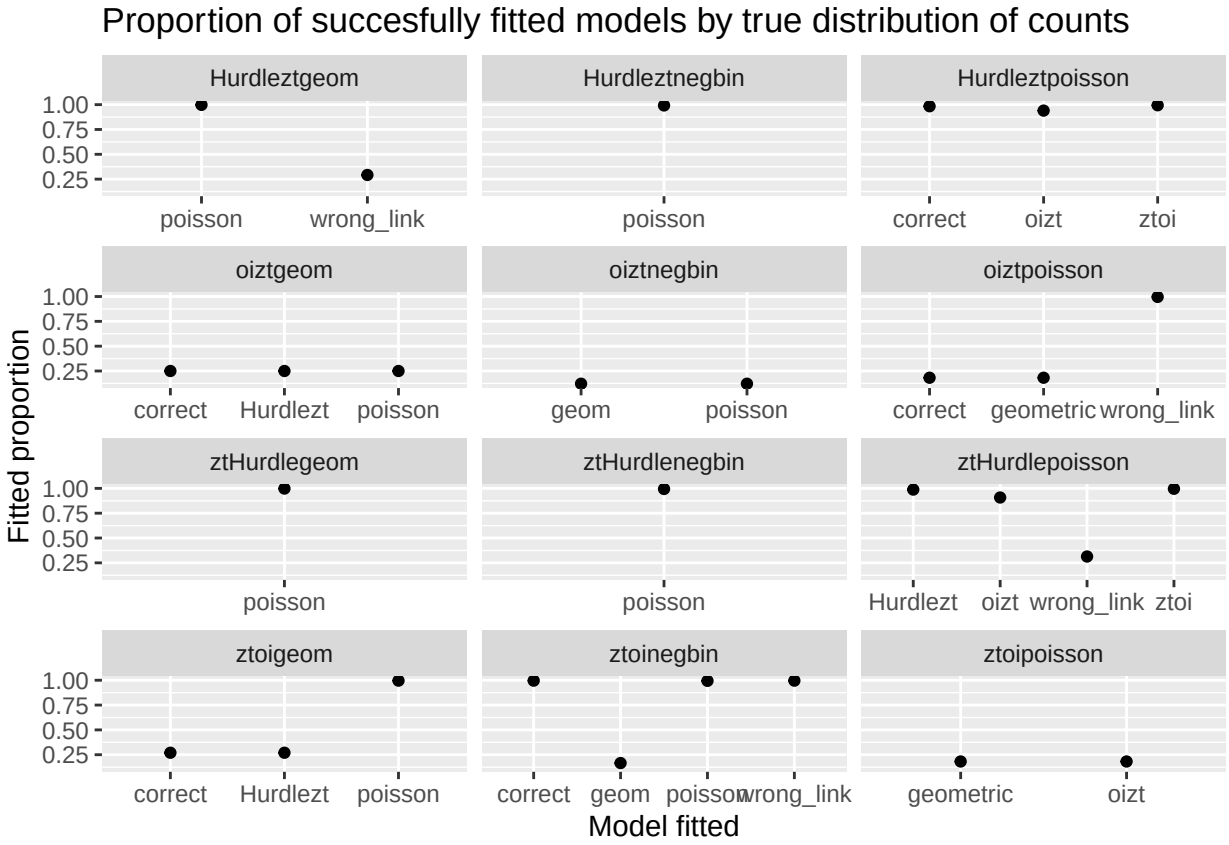
```
summarised_df <- results_data_frame |>
  group_by(data_generation, data_fitted) |>
  summarise(mean_point              = mean(point, na.rm = TRUE),
            mean_ci_length_norm     = mean(conf_int_normal_upper - conf_int_normal_lower, na.rm = TRUE)
            coverage_ci_norm        = mean((conf_int_normal_lower < 1000) & (1000 < conf_int_normal_upp
            mean_ci_length_log_norm = mean(conf_int_log_normal_upper - conf_int_log_normal_lower, na.rm
            coverage_ci_log_norm    = mean((conf_int_log_normal_lower < 1000) & (1000 < conf_int_log_no
            succesful_fits          = mean(!is.na(point)))
```

```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```

```
print(summarised_df, n=20)
```

```
## # A tibble: 80 x 8
## # Groups:   data_generation [12]
##    data_generation data_fitted mean_point mean_ci_length_norm coverage_ci_norm
##    <chr>           <chr>            <dbl>               <dbl>            <dbl>
##  1 Hurdleztgeom    correct        1.01e 3             1.66e 2            0.94
##  2 Hurdleztgeom    negbin         5.92e11             6.70e12            0.882
##  3 Hurdleztgeom    oizt           8.73e 2             7.00e 1            0
##  4 Hurdleztgeom    poisson        6.98e 2             2.38e 1            0
##  5 Hurdleztgeom    wrong_link     1.00e 3             1.68e 2            0.932
##  6 Hurdleztgeom    ztHurdle       1.09e 3             2.19e 2            0.692
##  7 Hurdleztgeom    ztoi           8.74e 2             7.93e 1            0
##  8 Hurdleztnegbin  correct        8.06e10             1.23e13            0.69
##  9 Hurdleztnegbin  geom           5.84e 2             6.16e 1            0
## 10 Hurdleztnegbin  oizt           6.17e 2             1.18e 2            0
## 11 Hurdleztnegbin  poisson        4.85e 2             7.58e 0            0
## 12 Hurdleztnegbin  wrong_link     6.51e14             3.72e14            0.696
## 13 Hurdleztnegbin  ztHurdle       2.99e10             5.75e 9            0.818
## 14 Hurdleztnegbin  ztoi           2.99e10             5.75e 9            0.818
## 15 Hurdleztpoisson correct        1.00e 3             1.03e 2            0.972
## 16 Hurdleztpoisson geometric      2.28e 3             8.20e 2            0
## 17 Hurdleztpoisson oizt           9.07e 2             4.23e 1            0
## 18 Hurdleztpoisson wrong_link     1.00e 3             1.03e 2            0.972
## 19 Hurdleztpoisson ztHurdle       1.06e 3             1.38e 2            0.704
## 20 Hurdleztpoisson ztoi           9.08e 2             4.23e 1            0
## # i 60 more rows
## # i 3 more variables: mean_ci_length_log_norm <dbl>,
## #   coverage_ci_log_norm <dbl>, succesful_fits <dbl>
```

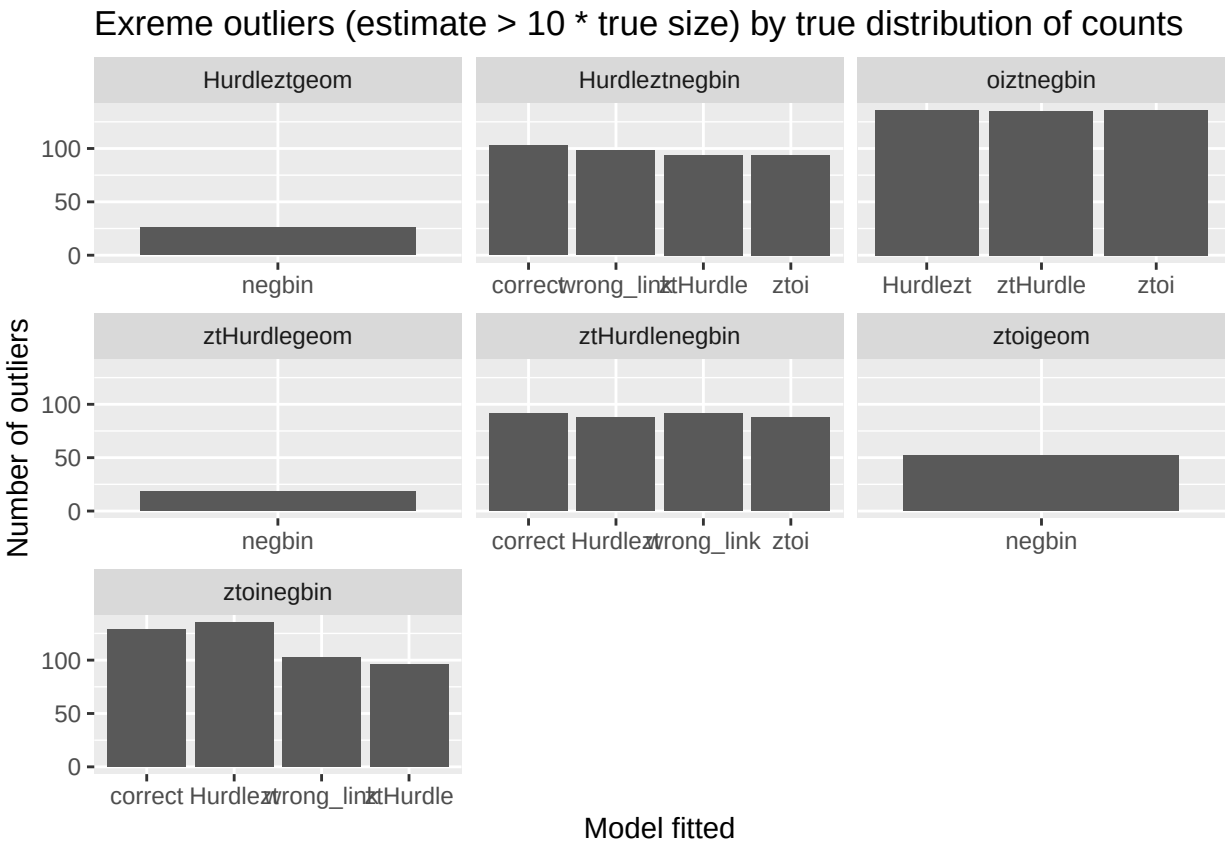
```
pp <- summarised_df |>
  subset(succesful_fits < 1) |>
  as.data.frame() |>
  mutate(data_generation = ordered(data_generation)) |>
  ggplot(aes(y = succesful_fits, x = data_fitted)) +
  geom_point() +
  facet_wrap(~data_generation, scales = c("free_x"), ncol = 3) +
  ylab("Fitted proportion") +
  xlab("Model fitted") +
  ggtitle("Proportion of succesfully fitted models by true distribution of counts")

pp
```



Visualising outliers (i.e. when estimated regression parameters tend to boundary):

```
results_data_frame |>
  subset(!is.na(point)) |>
  subset(point > 10000) |>
  group_by(data_generation, data_fitted) |>
  summarise(n = n()) |>
  ggplot(aes(x = data_fitted, weight = n)) +
  geom_bar() +
  facet_wrap(~ data_generation, scales = c("free_x")) +
  ylab("Number of outliers") +
  xlab("Model fitted") +
  ggtitle("Exreme outliers (estimate > 10 * true size) by true distribution of counts")
```

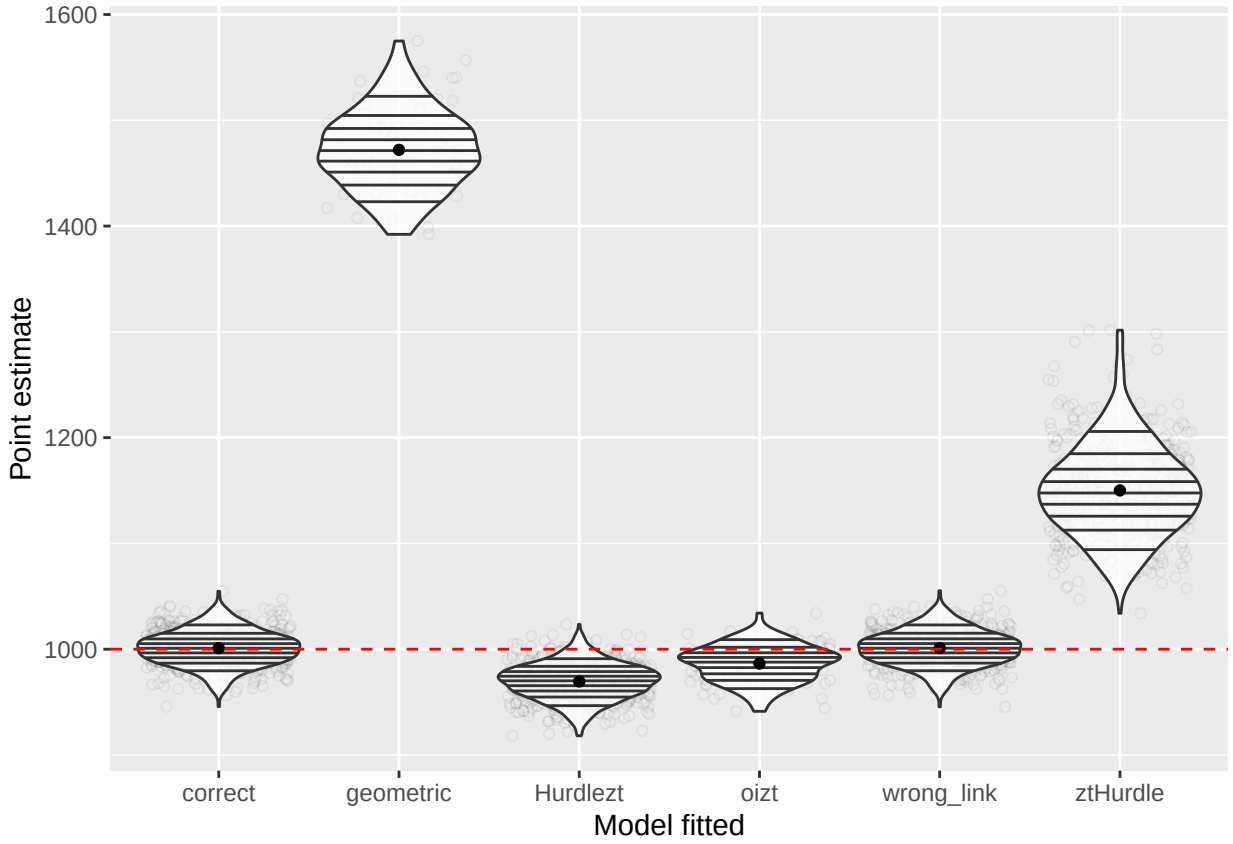
```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```



Point estimates

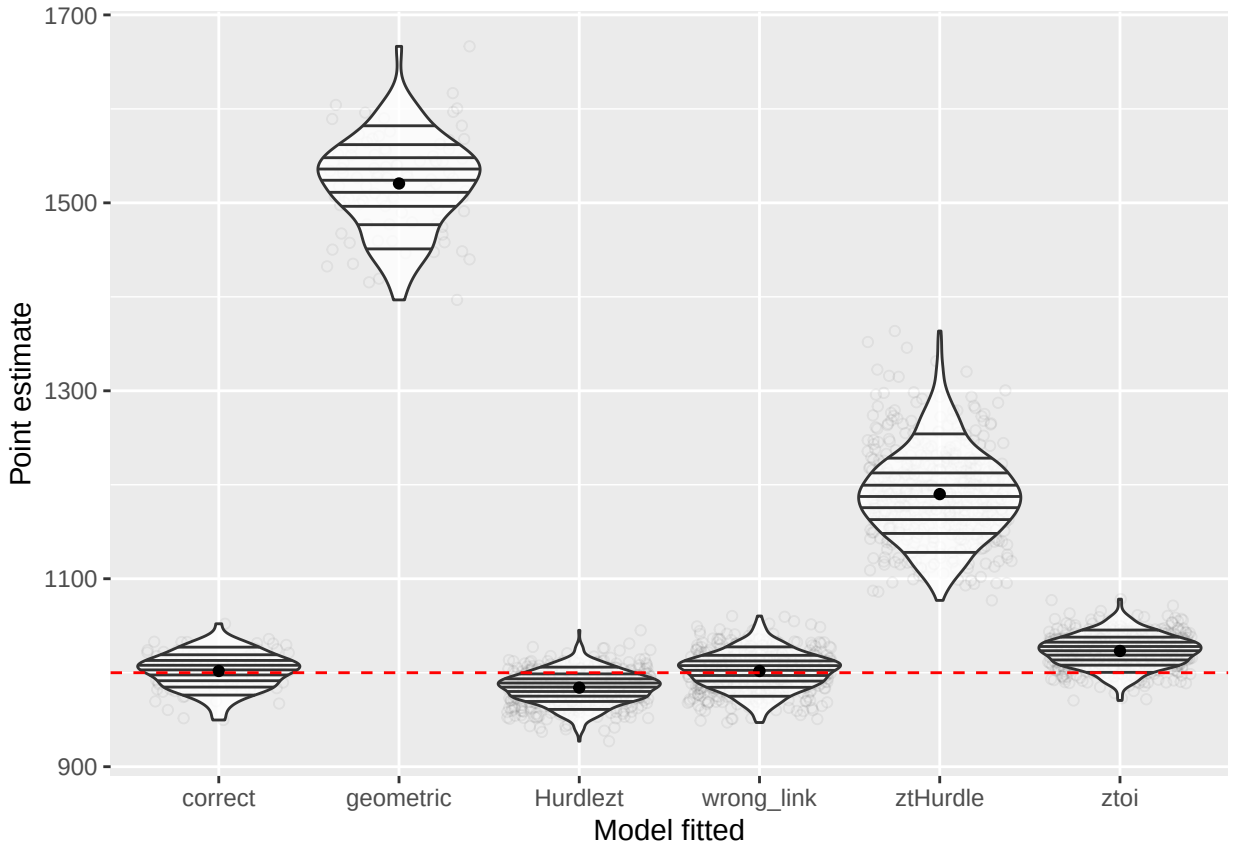
Results for counts generated by ztoipoisson:

```
p1 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztoipoisson")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p1
```



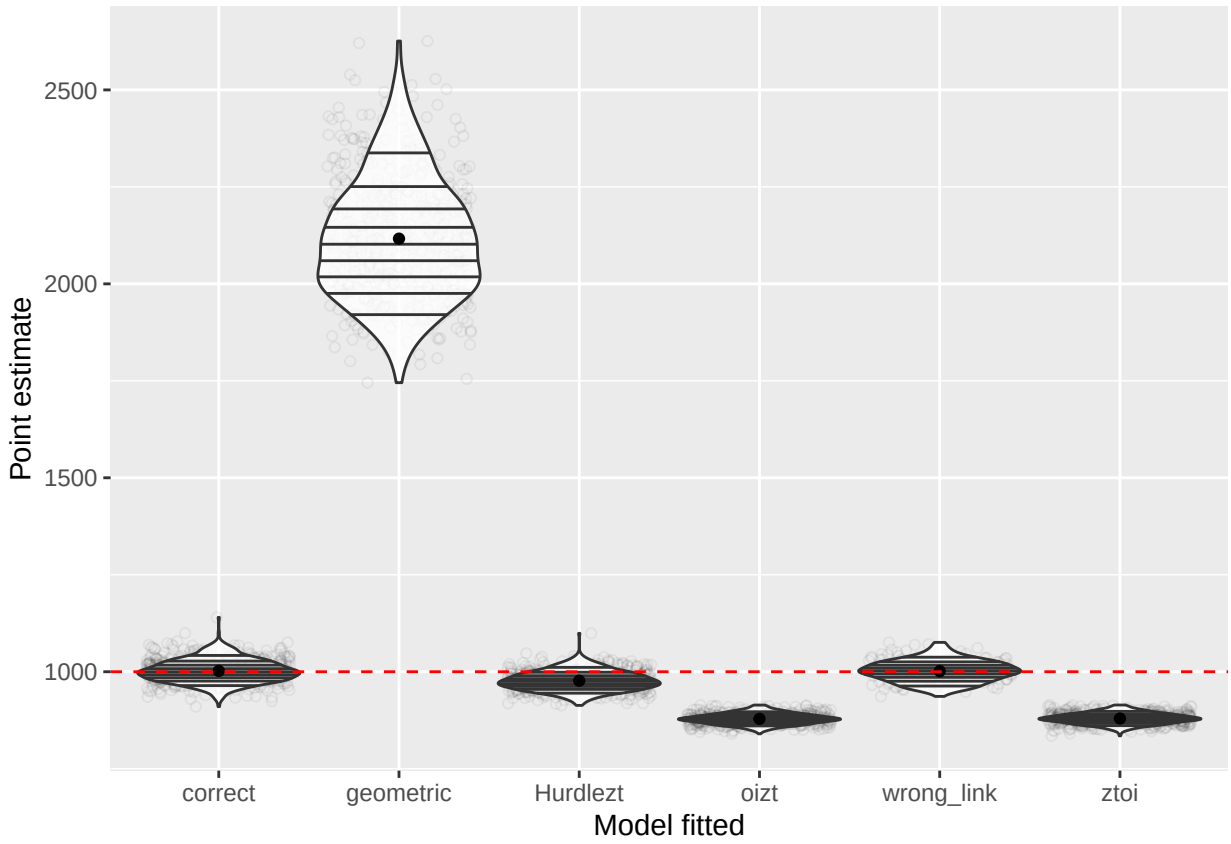
Results for counts generated by oiztpoisson:

```
p2 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "oiztpoisson")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p2
```



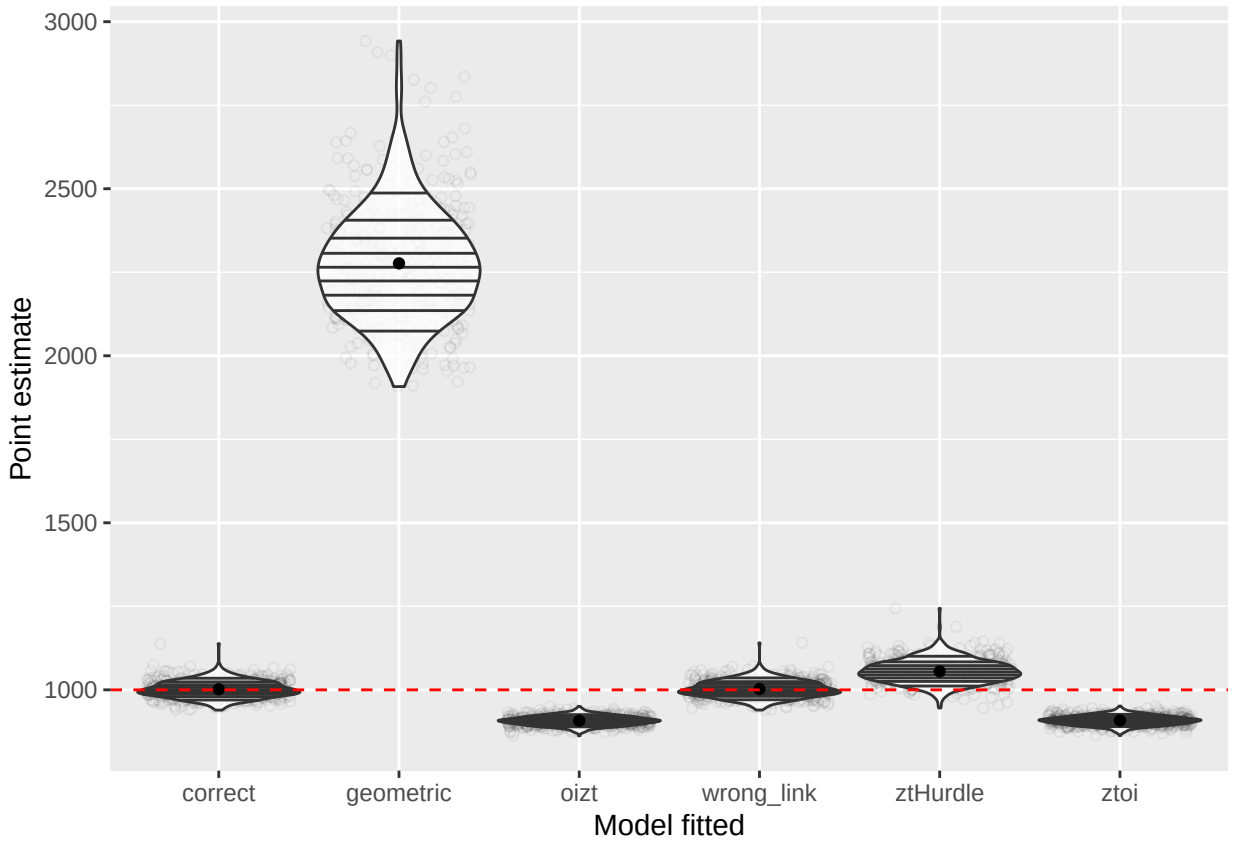
Results for counts generated by ztHurdlepoisson:

```
p3 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztHurdlepoisson")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p3
```



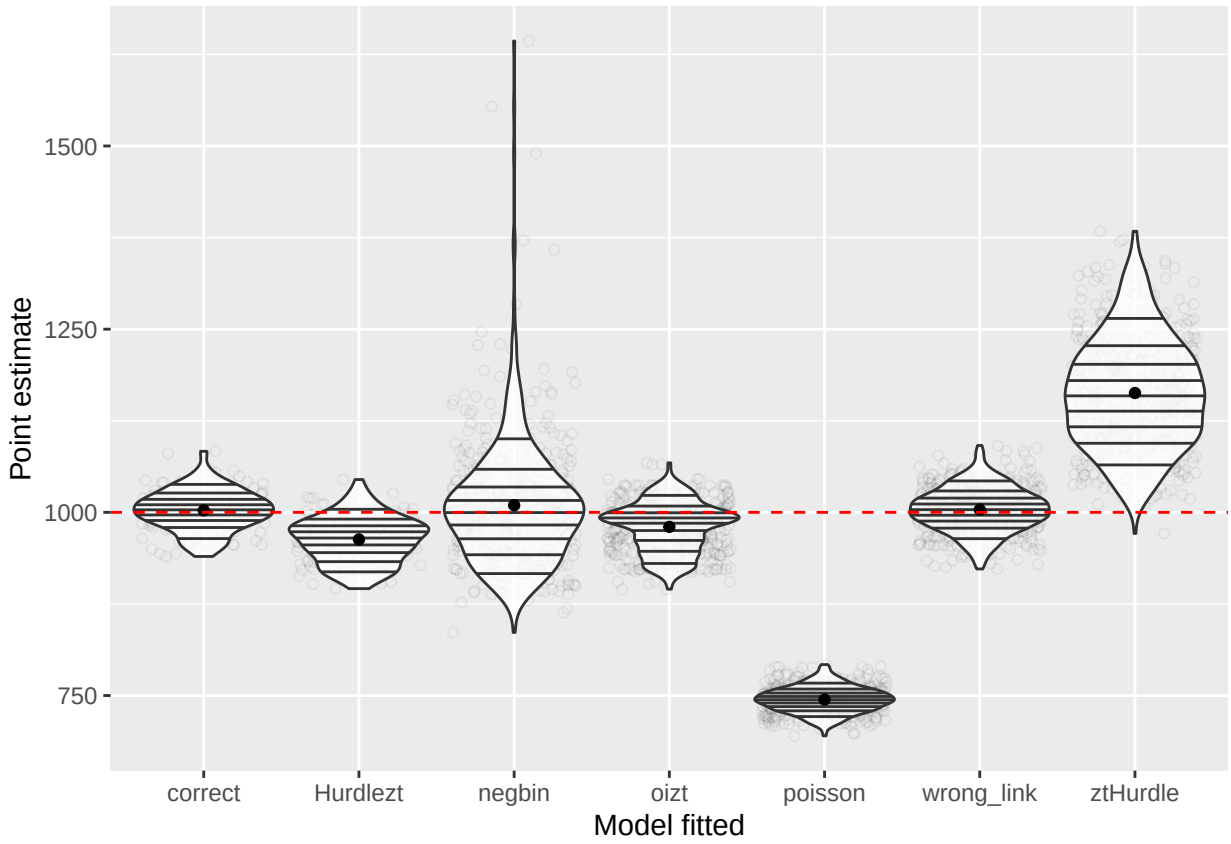
Results for counts generated by hurdleztpoisson:

```
p4 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "Hurdleztpoisson")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p4
```



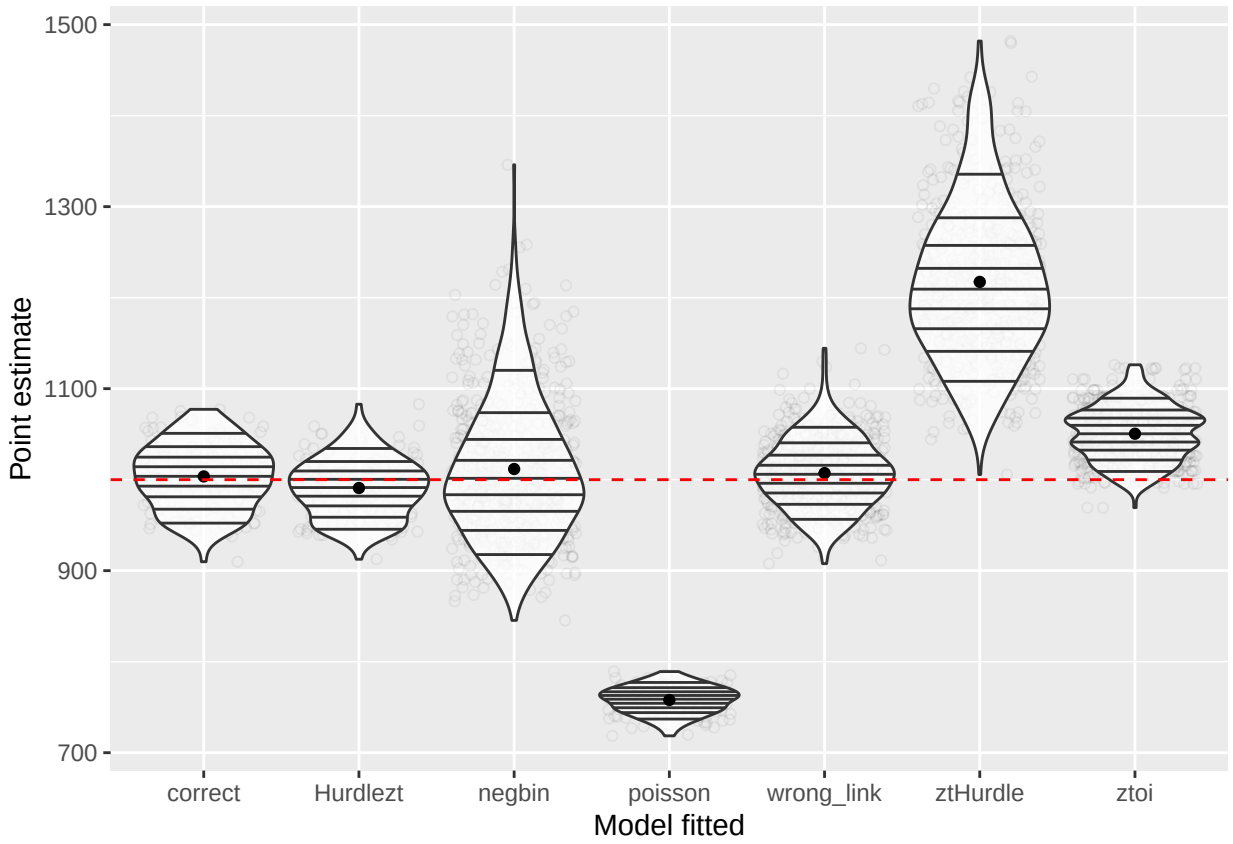
Results for counts generated by ztoigeom:

```
p5 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztoigeom")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p5
```



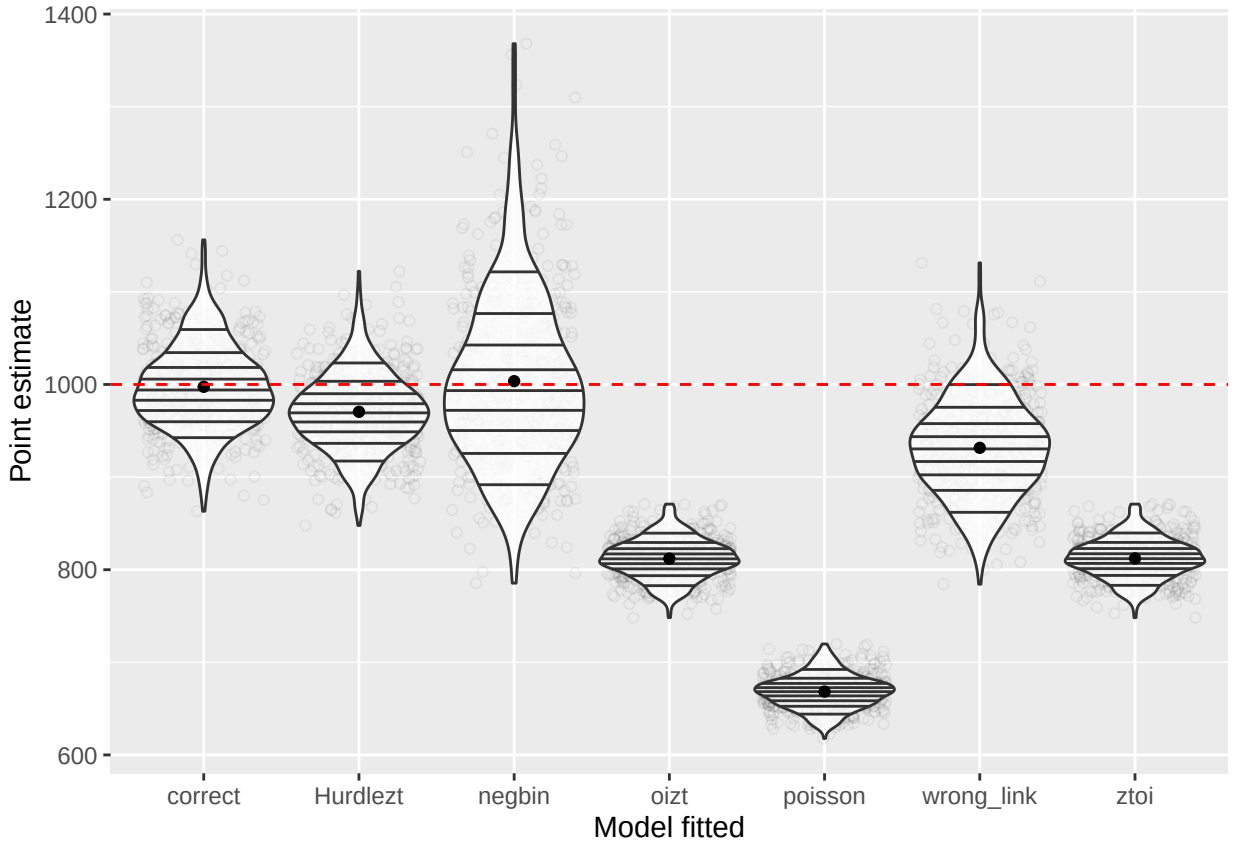
Results for counts generated by oiztgeom:

```
p6 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "oiztgeom")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p6
```



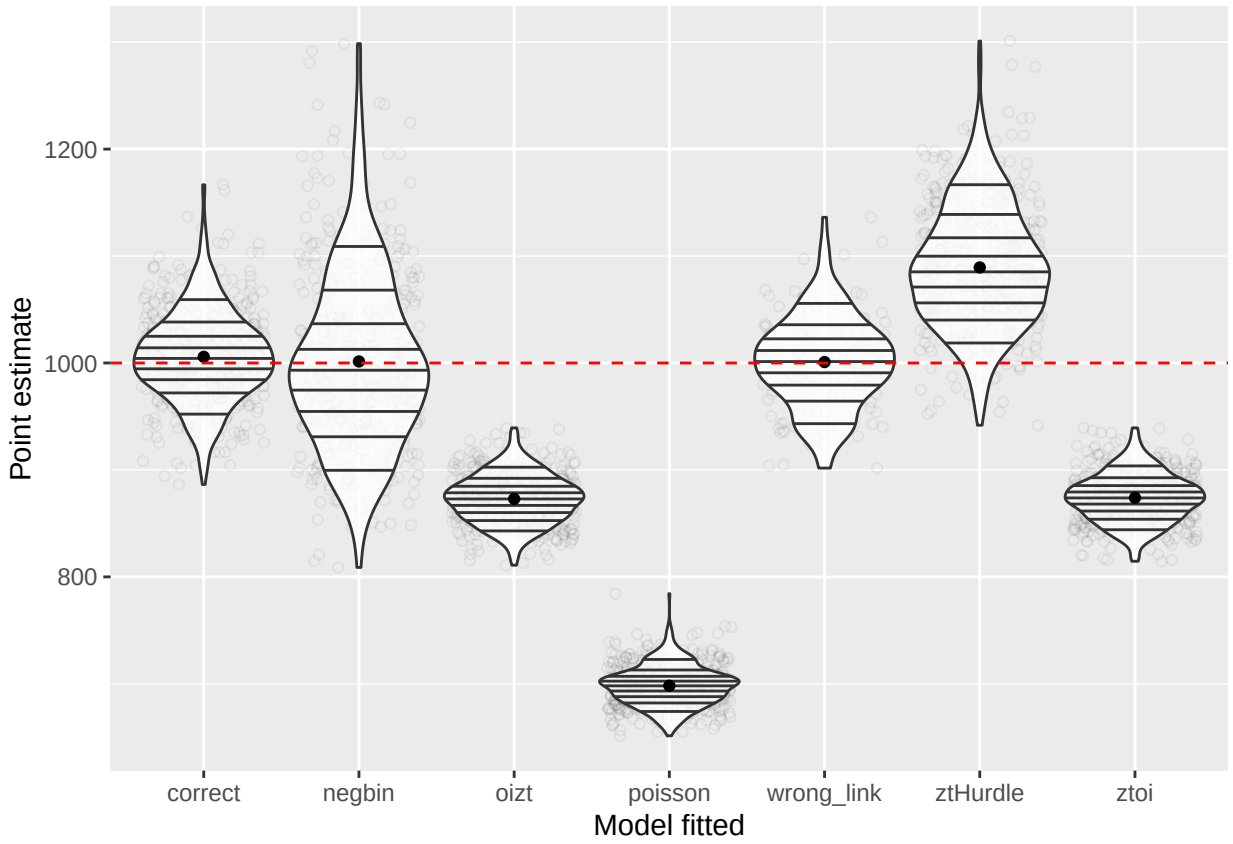
Results for counts generated by ztHurdlegeom:

```
p7 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztHurdlegeom")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p7
```



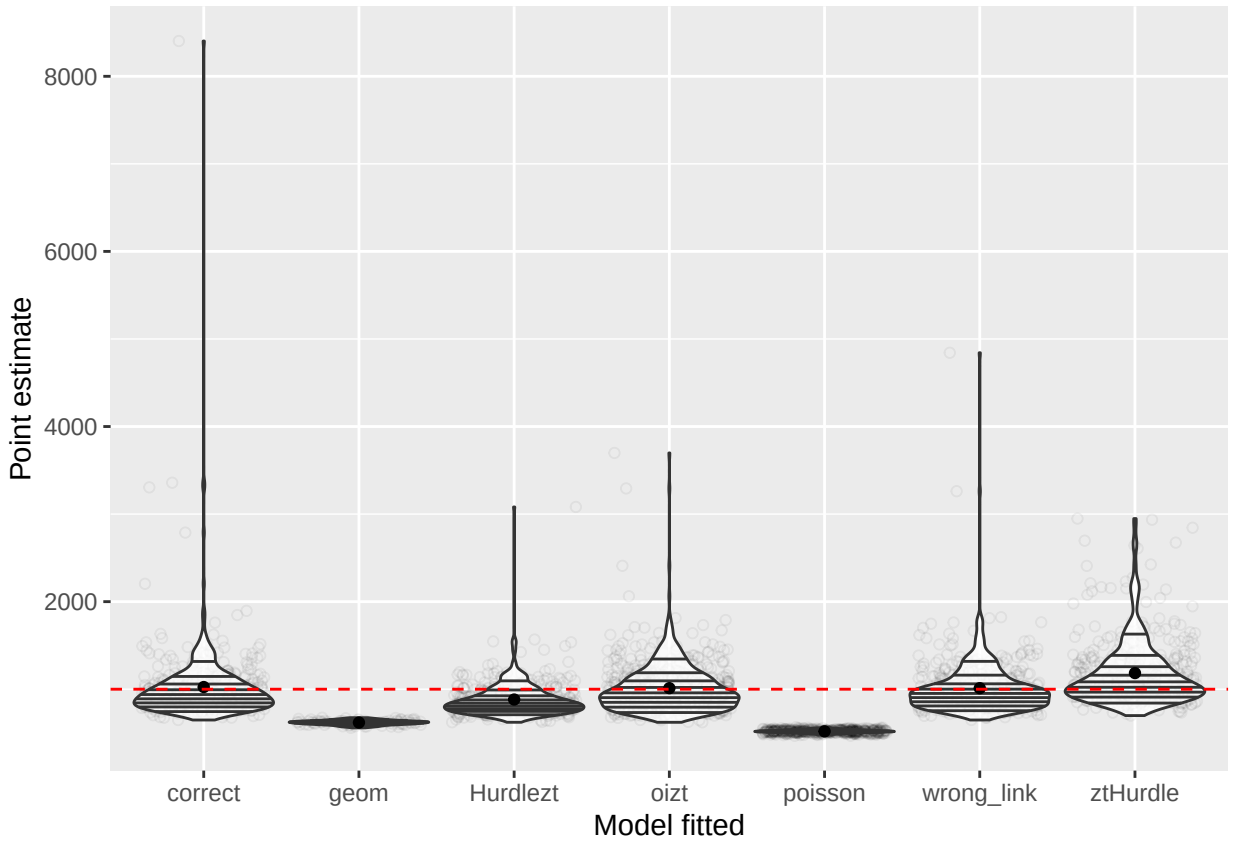
Results for counts generated by hurdleztgeom:

```
p8 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "Hurdleztgeom")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p8
```



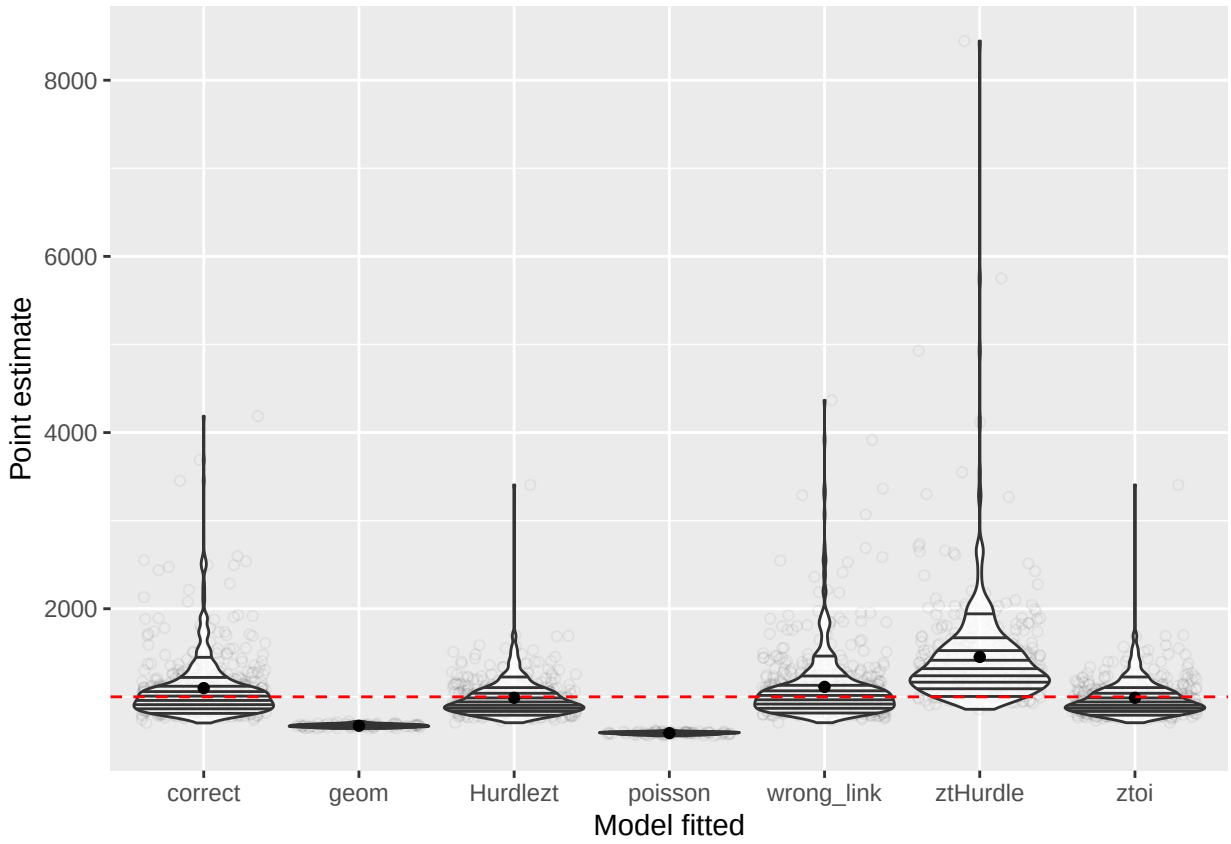
Results for counts generated by ztoinegbin:

```
p9 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztoinegbin")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p9
```



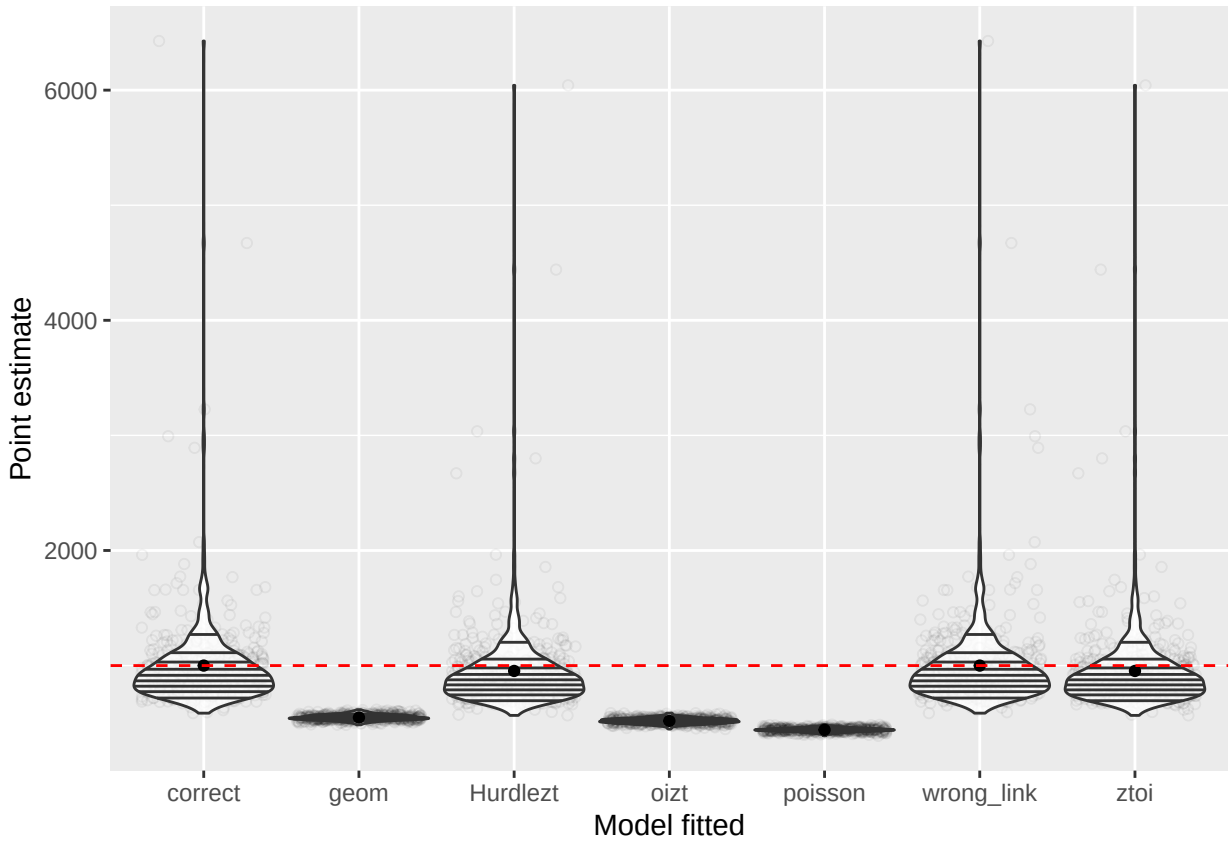
Results for counts generated by oiztnegbin:

```
p10 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "oiztnegbin")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p10
```



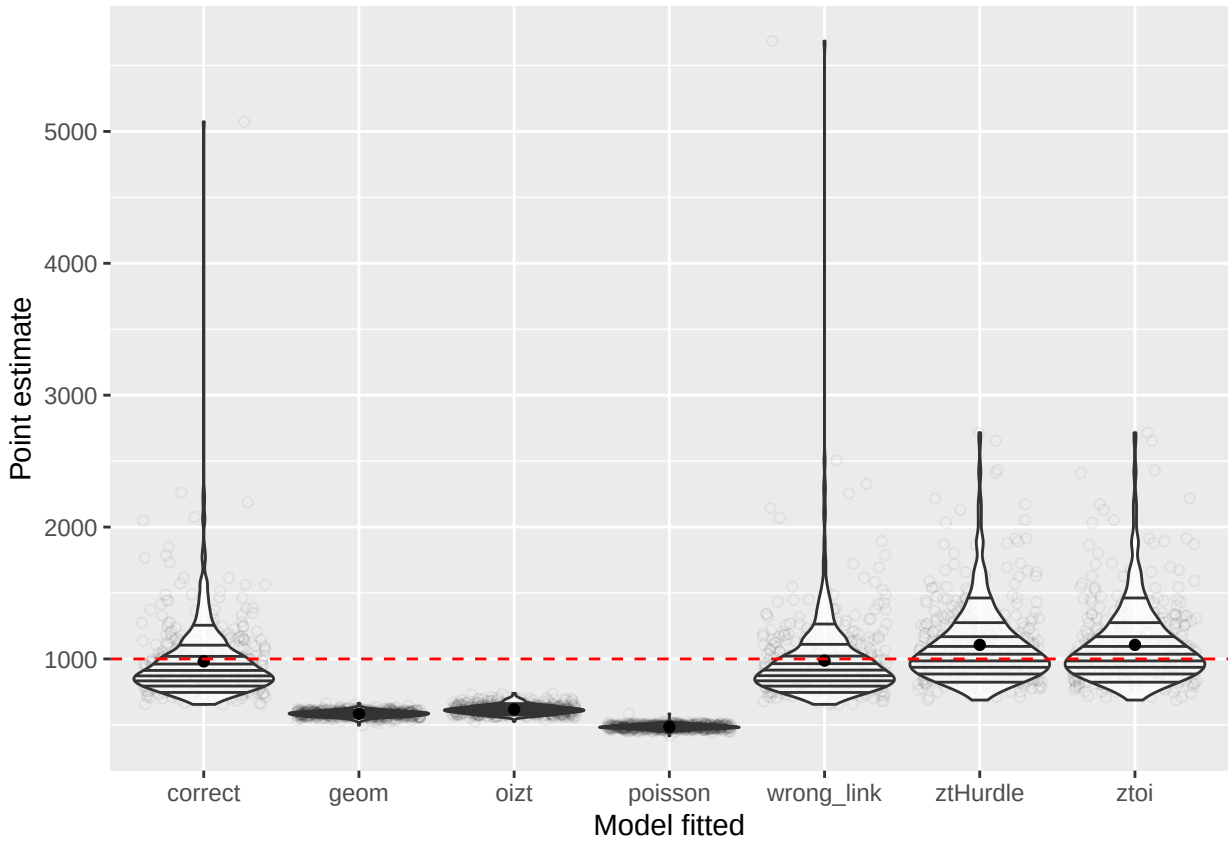
Results for counts generated by ztHurdlenegbin:

```
p11 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "ztHurdlenegbin")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p11
```



Results for counts generated by hurdleztnegbin:

```
p12 <- results_data_frame |>
  subset(!is.na(point) & (data_generation == "Hurdleztnegbin")) |>
  subset(point < 25000) |>
  ggplot(aes(x = data_fitted, y = point)) +
  geom_jitter(alpha = 0.05, shape = 1) +
  geom_violin(alpha = 0.8, draw_quantiles = 1:9 / 10, scale = "width") +
  stat_summary(fun = function(x) mean(x, na.rm = TRUE), geom = "point") +
  geom_hline(yintercept = 1000, linetype="dashed", color = "red") +
  ylab("Point estimate") +
  xlab("Model fitted")

p12
```



Confidence intervals

Normal

Exact binomial tests for coverage of lognormal confindence intervals with $H_0 : p = 0.95, H_1 = \neg H_1$:

```
dd <- results_data_frame |>
  subset(!is.na(point)) |>
  subset(point < 25000) |>
  mutate(covr_norm = (conf_int_normal_lower < 1000) & (conf_int_normal_upper > 1000),
         covr_log  = (conf_int_log_normal_lower < 1000) & (conf_int_log_normal_upper > 1000)) |>
  group_by(data_generation, data_fitted) |>
  summarise(n = n(),
            mean = mean(covr_norm, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```

```
dd <- cbind(dd, p_value = NA, lower = NA, upper = NA)

for (x in 1:NROW(dd)) {
  jj <- binom.test(x = as.numeric(dd[x, 4]) * as.integer(dd[x, 3]), n = as.integer(dd[x, 3]), p = .95)
  # this jj object has some very weird interactions with the rest of R ecosystem
  dd[x, 5] <- jj$p.value |> as.numeric()
  dd[x, 6] <- jj[[4]][1]
```

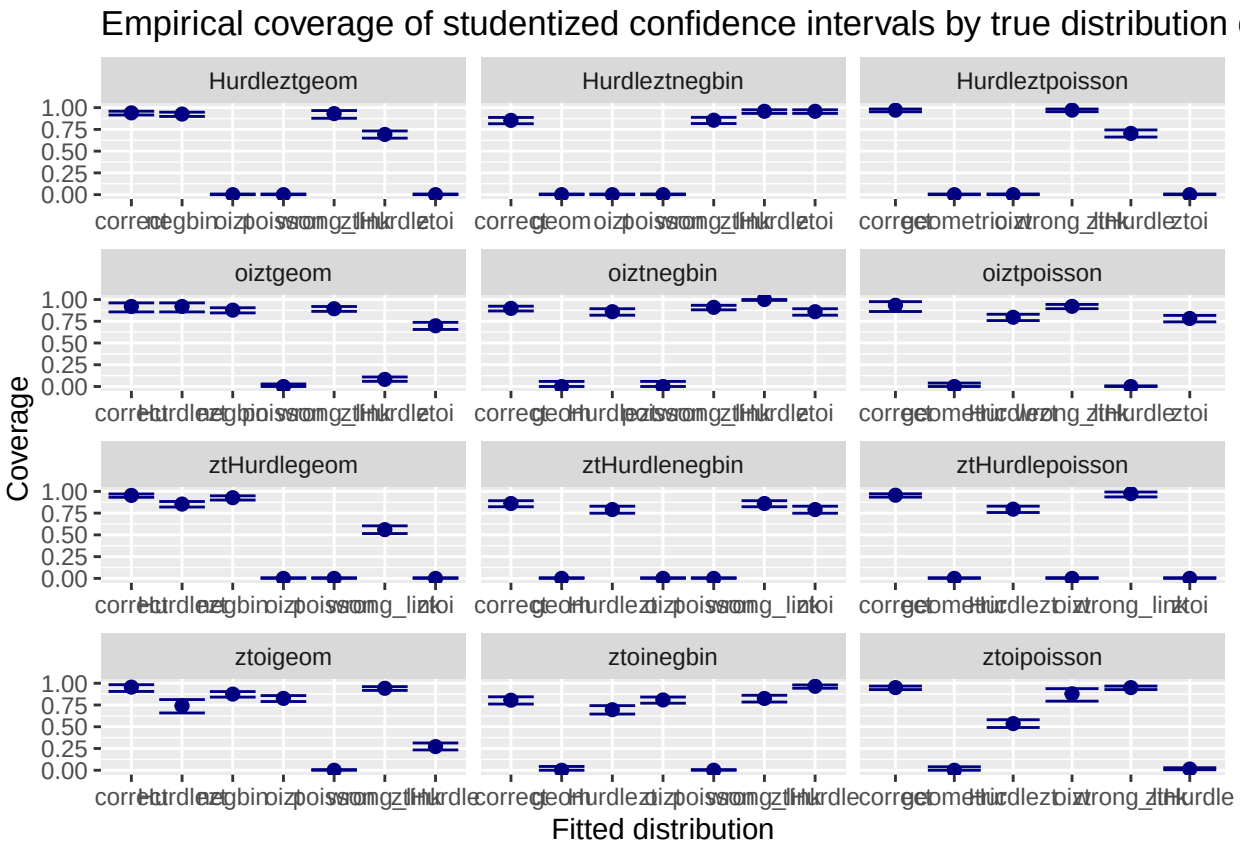
```
  dd[x, 7] <- jj[[4]][2]
}

print(dd[, c(1:2, 4, 5)] |> mutate(p_value = round(p_value, digits = 4)),
      n = NROW(dd))
```

```
## # A tibble: 80 x 4
## # Groups:   data_generation [12]
##    data_generation data_fitted  mean p_value
##    <chr>           <chr>       <dbl>   <dbl>
##  1 Hurdleztgeom    correct     0.94   0.304
##  2 Hurdleztgeom    negbin      0.926  0.026
##  3 Hurdleztgeom    oizt        0      0
##  4 Hurdleztgeom    poisson     0      0
##  5 Hurdleztgeom    wrong_link  0.932  0.337
##  6 Hurdleztgeom    ztHurdle    0.692  0
##  7 Hurdleztgeom    ztoi        0      0
##  8 Hurdleztnegbin  correct     0.854  0
##  9 Hurdleztnegbin  geom        0      0
## 10 Hurdleztnegbin  oizt        0      0
## 11 Hurdleztnegbin  poisson     0      0
## 12 Hurdleztnegbin  wrong_link  0.856  0
## 13 Hurdleztnegbin  ztHurdle    0.958  0.568
## 14 Hurdleztnegbin  ztoi        0.958  0.568
## 15 Hurdleztpoisson correct     0.972  0.029
## 16 Hurdleztpoisson geometric   0      0
## 17 Hurdleztpoisson oizt        0      0
## 18 Hurdleztpoisson wrong_link  0.972  0.0232
## 19 Hurdleztpoisson ztHurdle    0.704  0
## 20 Hurdleztpoisson ztoi        0      0
## 21 oiztgeom        Hurdlezt    0.92   0.144
## 22 oiztgeom        correct     0.92   0.144
## 23 oiztgeom        negbin      0.878  0
## 24 oiztgeom        poisson     0      0
## 25 oiztgeom        wrong_link  0.894  0
## 26 oiztgeom        ztHurdle    0.082  0
## 27 oiztgeom        ztoi        0.698  0
## 28 oiztnegbin      Hurdlezt    0.860  0
## 29 oiztnegbin      correct     0.898  0
## 30 oiztnegbin      geom        0      0
## 31 oiztnegbin      poisson     0      0
## 32 oiztnegbin      wrong_link  0.91   0.0002
## 33 oiztnegbin      ztHurdle    1      0
## 34 oiztnegbin      ztoi        0.860  0
## 35 oiztpoisson     Hurdlezt    0.796  0
## 36 oiztpoisson     correct     0.934  0.464
## 37 oiztpoisson     geometric   0      0
## 38 oiztpoisson     wrong_link  0.922  0.0071
## 39 oiztpoisson     ztHurdle    0      0
## 40 oiztpoisson     ztoi        0.782  0
## 41 ztHurdlegeom    Hurdlezt    0.854  0
## 42 ztHurdlegeom    correct     0.954  0.758
## 43 ztHurdlegeom    negbin      0.927  0.0277
```

```
## 44 ztHurdlegeom    oizt        0      0
## 45 ztHurdlegeom    poisson     0      0
## 46 ztHurdlegeom    wrong_link  0.56   0
## 47 ztHurdlegeom    ztoi        0      0
## 48 ztHurdlenegbin  Hurdlezt    0.791  0
## 49 ztHurdlenegbin  correct     0.861  0
## 50 ztHurdlenegbin  geom        0      0
## 51 ztHurdlenegbin  oizt        0      0
## 52 ztHurdlenegbin  poisson     0      0
## 53 ztHurdlenegbin  wrong_link  0.861  0
## 54 ztHurdlenegbin  ztoi        0.791  0
## 55 ztHurdlepoisson Hurdlezt    0.796  0
## 56 ztHurdlepoisson correct     0.956  0.608
## 57 ztHurdlepoisson geometric   0      0
## 58 ztHurdlepoisson oizt        0      0
## 59 ztHurdlepoisson wrong_link  0.975  0.198
## 60 ztHurdlepoisson ztoi        0      0
## 61 ztoigeom        Hurdlezt    0.741  0
## 62 ztoigeom        correct     0.956  1
## 63 ztoigeom        negbin      0.875  0
## 64 ztoigeom        oizt        0.826  0
## 65 ztoigeom        poisson     0      0
## 66 ztoigeom        wrong_link  0.942  0.410
## 67 ztoigeom        ztHurdle    0.272  0
## 68 ztoinegbin      Hurdlezt    0.696  0
## 69 ztoinegbin      correct     0.805  0
## 70 ztoinegbin      geom        0      0
## 71 ztoinegbin      oizt        0.808  0
## 72 ztoinegbin      poisson     0      0
## 73 ztoinegbin      wrong_link  0.825  0
## 74 ztoinegbin      ztHurdle    0.965  0.171
## 75 ztoipoisson     Hurdlezt    0.536  0
## 76 ztoipoisson     correct     0.95   1
## 77 ztoipoisson     geometric   0      0
## 78 ztoipoisson     oizt        0.879  0.0058
## 79 ztoipoisson     wrong_link  0.95   1
## 80 ztoipoisson     ztHurdle    0.014  0
```

Visual results with confidence intervals:

```
qq1 <- dd |>
  ggplot(aes(x = data_fitted)) +
  facet_wrap(~ data_generation, scales = c("free_x"), ncol = 3) +
  geom_point(aes(y = mean), colour = "navy", cex = 2) +
  geom_errorbar(aes(ymin = lower, ymax = upper), colour = "navy") +
  ggtitle("Empirical coverage of studentized confidence intervals by true distribution of counts") +
  xlab("Fitted distribution") +
  ylab("Coverage")

qq1
```

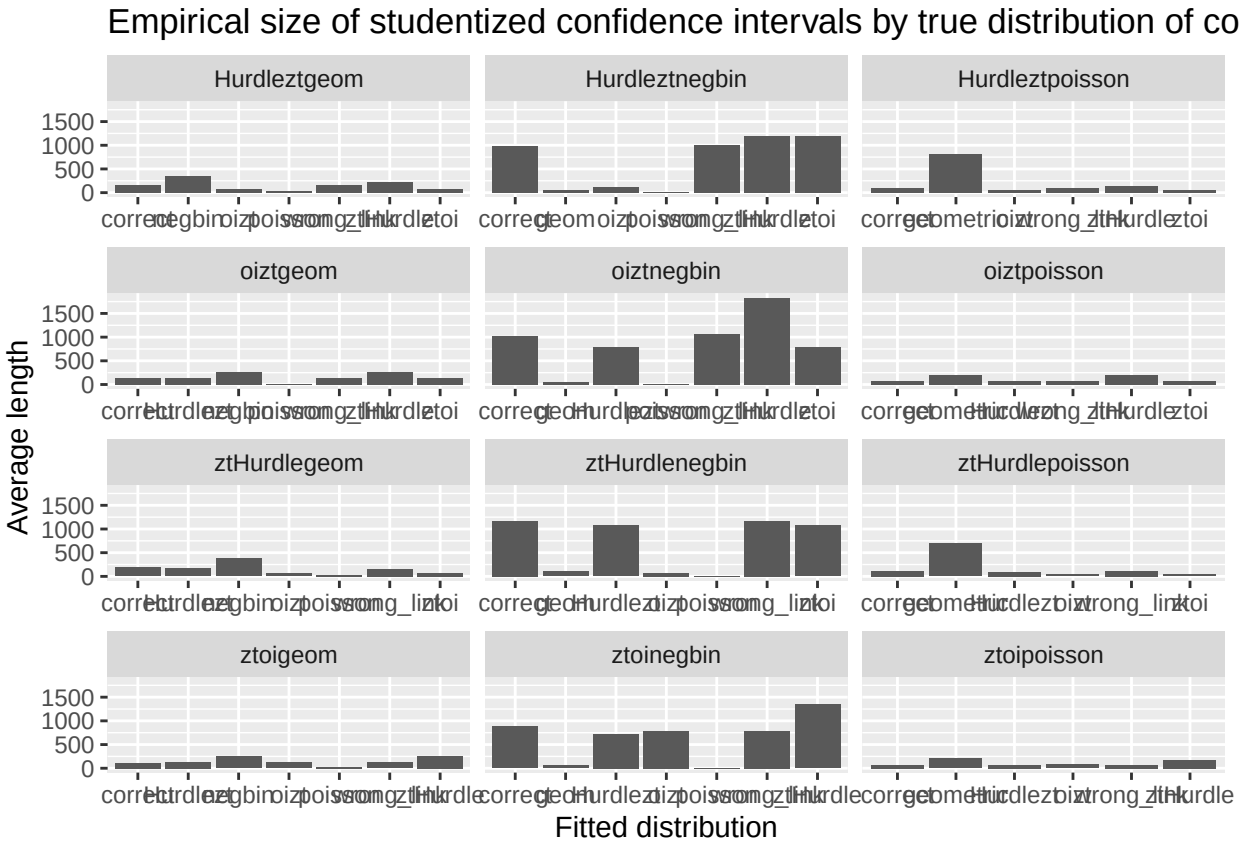


Average sizes of confidence intervals:

```
qq2 <- results_data_frame |>
  subset(!is.na(point)) |>
  subset(point < 25000) |>
  group_by(data_generation, data_fitted) |>
  summarise(len = mean(conf_int_normal_upper - conf_int_normal_lower, na.rm = TRUE)) |>
  ggplot(aes(x = data_fitted, weight = len)) +
  geom_bar() +
  facet_wrap(~ data_generation, scales = c("free_x"), ncol = 3) +
  ylab("Average length") +
  xlab("Fitted distribution") +
  ggtitle("Empirical size of studentized confidence intervals by true distribution of counts")
```

```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```

```
qq2
```



Logormal

Exact binomial tests for coverage of normal confindence intervals with $H_0 : p = 0.95, H_1 = \neg H_1$:

```
dd <- results_data_frame |>
  subset(!is.na(point)) |>
  subset(point < 25000) |>
  mutate(covr_norm = (conf_int_normal_lower < 1000) & (conf_int_normal_upper > 1000),
         covr_log  = (conf_int_log_normal_lower < 1000) & (conf_int_log_normal_upper > 1000)) |>
  group_by(data_generation, data_fitted) |>
  summarise(n = n(),
            mean = mean(covr_log, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```

```
dd <- cbind(dd, p_value = NA, lower = NA, upper = NA)

for (x in 1:NROW(dd)) {
  jj <- binom.test(x = as.numeric(dd[x, 4]) * as.integer(dd[x, 3]), n = as.integer(dd[x, 3]), p = .95)
  # this jj object has some very weird interactions with the rest of R ecosystem
  dd[x, 5] <- jj$p.value |> as.numeric()
  dd[x, 6] <- jj[[4]][1]
  dd[x, 7] <- jj[[4]][2]
}
```

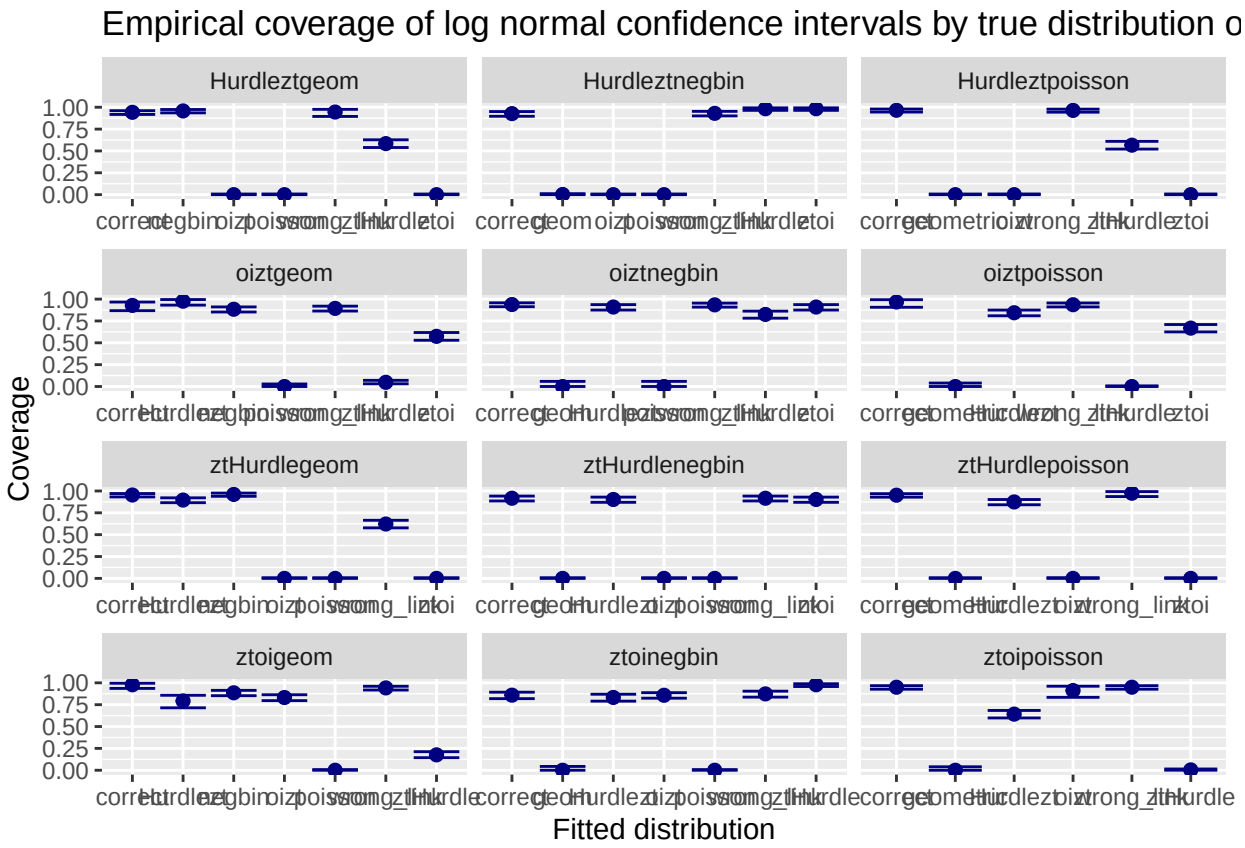
```r
print(dd[, c(1:2, 4, 5)] |> mutate(p_value = round(p_value, digits = 4)),
      n = NROW(dd))
```

```
## # A tibble: 80 x 4
## # Groups:   data_generation [12]
##    data_generation data_fitted  mean p_value
##    <chr>           <chr>       <dbl>   <dbl>
##  1 Hurdleztgeom    correct     0.942  0.410
##  2 Hurdleztgeom    negbin      0.958  0.526
##  3 Hurdleztgeom    oizt        0      0
##  4 Hurdleztgeom    poisson     0      0
##  5 Hurdleztgeom    wrong_link  0.945  0.704
##  6 Hurdleztgeom    ztHurdle    0.584  0
##  7 Hurdleztgeom    ztoi        0      0
##  8 Hurdleztnegbin  correct     0.927  0.0488
##  9 Hurdleztnegbin  geom        0.002  0
## 10 Hurdleztnegbin  oizt        0      0
## 11 Hurdleztnegbin  poisson     0      0
## 12 Hurdleztnegbin  wrong_link  0.930  0.0844
## 13 Hurdleztnegbin  ztHurdle    0.983  0.0009
## 14 Hurdleztnegbin  ztoi        0.983  0.0009
## 15 Hurdleztpoisson correct     0.965  0.121
## 16 Hurdleztpoisson geometric   0      0
## 17 Hurdleztpoisson oizt        0      0
## 18 Hurdleztpoisson wrong_link  0.964  0.181
## 19 Hurdleztpoisson ztHurdle    0.566  0
## 20 Hurdleztpoisson ztoi        0      0
## 21 oiztgeom        Hurdlezt    0.976  0.220
## 22 oiztgeom        correct     0.928  0.298
## 23 oiztgeom        negbin      0.884  0
## 24 oiztgeom        poisson     0      0
## 25 oiztgeom        wrong_link  0.894  0
## 26 oiztgeom        ztHurdle    0.048  0
## 27 oiztgeom        ztoi        0.574  0
## 28 oiztnegbin      Hurdlezt    0.909  0.001
## 29 oiztnegbin      correct     0.938  0.217
## 30 oiztnegbin      geom        0      0
## 31 oiztnegbin      poisson     0      0
## 32 oiztnegbin      wrong_link  0.934  0.101
## 33 oiztnegbin      ztHurdle    0.825  0
## 34 oiztnegbin      ztoi        0.909  0.001
## 35 oiztpoisson     Hurdlezt    0.844  0
## 36 oiztpoisson     correct     0.967  0.631
## 37 oiztpoisson     geometric   0      0
## 38 oiztpoisson     wrong_link  0.936  0.149
## 39 oiztpoisson     ztHurdle    0      0
## 40 oiztpoisson     ztoi        0.668  0
## 41 ztHurdlegeom    Hurdlezt    0.896  0
## 42 ztHurdlegeom    correct     0.954  0.758
## 43 ztHurdlegeom    negbin      0.961  0.346
## 44 ztHurdlegeom    oizt        0      0
## 45 ztHurdlegeom    poisson     0      0
## 46 ztHurdlegeom    wrong_link  0.622  0
```

```
## 47 ztHurdlegeom    ztoi        0      0
## 48 ztHurdlenegbin  Hurdlezt    0.903  0.0001
## 49 ztHurdlenegbin  correct     0.917  0.0042
## 50 ztHurdlenegbin  geom        0      0
## 51 ztHurdlenegbin  oizt        0      0
## 52 ztHurdlenegbin  poisson     0      0
## 53 ztHurdlenegbin  wrong_link  0.917  0.0042
## 54 ztHurdlenegbin  ztoi        0.903  0.0001
## 55 ztHurdlepoisson Hurdlezt    0.874  0
## 56 ztHurdlepoisson correct     0.952  0.918
## 57 ztHurdlepoisson geometric   0      0
## 58 ztHurdlepoisson oizt        0      0
## 59 ztHurdlepoisson wrong_link  0.975  0.198
## 60 ztHurdlepoisson ztoi        0      0
## 61 ztoigeom        Hurdlezt    0.793  0
## 62 ztoigeom        correct     0.978  0.166
## 63 ztoigeom        negbin      0.886  0
## 64 ztoigeom        oizt        0.832  0
## 65 ztoigeom        poisson     0      0
## 66 ztoigeom        wrong_link  0.942  0.410
## 67 ztoigeom        ztHurdle    0.176  0
## 68 ztoinegbin      Hurdlezt    0.833  0
## 69 ztoinegbin      correct     0.859  0
## 70 ztoinegbin      geom        0      0
## 71 ztoinegbin      oizt        0.858  0
## 72 ztoinegbin      poisson     0      0
## 73 ztoinegbin      wrong_link  0.873  0
## 74 ztoinegbin      ztHurdle    0.978  0.0081
## 75 ztoipoisson     Hurdlezt    0.642  0
## 76 ztoipoisson     correct     0.95   1
## 77 ztoipoisson     geometric   0      0
## 78 ztoipoisson     oizt        0.912  0.140
## 79 ztoipoisson     wrong_link  0.95   1
## 80 ztoipoisson     ztHurdle    0.004  0
```

Visual results with confidence intervals:

```
qq3 <- dd |>
  ggplot(aes(x = data_fitted)) +
  facet_wrap(~ data_generation, scales = c("free_x"), ncol = 3) +
  geom_point(aes(y = mean), colour = "navy", cex = 2) +
  geom_errorbar(aes(ymin = lower, ymax = upper), colour = "navy") +
  ggtitle("Empirical coverage of log normal confidence intervals by true distribution of counts") +
  xlab("Fitted distribution") +
  ylab("Coverage")

qq3
```

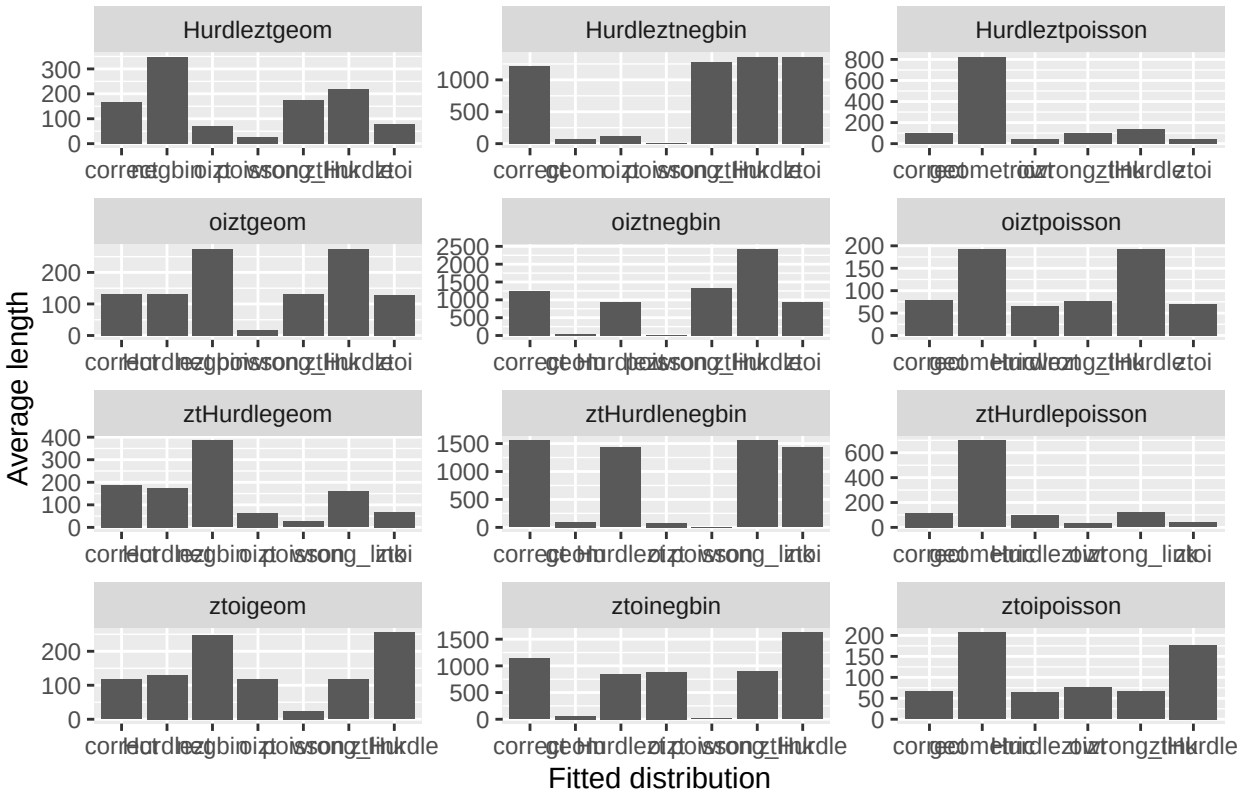


Average sizes of confidence intervals:

```
qq4 <- results_data_frame |>
  subset(!is.na(point)) |>
  subset(point < 25000) |>
  group_by(data_generation, data_fitted) |>
  summarise(len = mean(conf_int_log_normal_upper - conf_int_log_normal_lower, na.rm = TRUE)) |>
  ggplot(aes(x = data_fitted, weight = len)) +
  geom_bar() +
  facet_wrap(~ data_generation, scales = "free", ncol = 3) +
  ylab("Average length") +
  xlab("Fitted distribution") +
  ggtitle("Empirical size of log normal confidence intervals by true distribution of counts")
```

```
## `summarise()` has grouped output by 'data_generation'. You can override using
## the `.groups` argument.
```

```
qq4
```


Empirical size of log normal confidence intervals by true distribution of cour
Hurdleztgeom
Hurdleztnegbin
Hurdleztpoisson
oiztgeom
oiztnegbin
oiztpoisson
ztHurdlegeom
ztHurdlenegbin
ztHurdlepoisson
ztoigeom
ztoinegbin
ztoipoisson
Average length
Fitted distribution