Goodness of fit

```
library(knitr)
library(dplyr)
library(ggplot2)

load("data/goodness_of_fit_dc.RData")
```

```
kable(tab_results[1:4,], digits = 4)
```

	Home	Draw	Away
Observed	0.4882	0.2686	0.2433
Model 0 (min 0)	0.4827	0.2608	0.2565
Model 3 (min 0)	0.4846	0.2788	0.2366
Model 8 (min 0)	0.4853	0.2765	0.2382

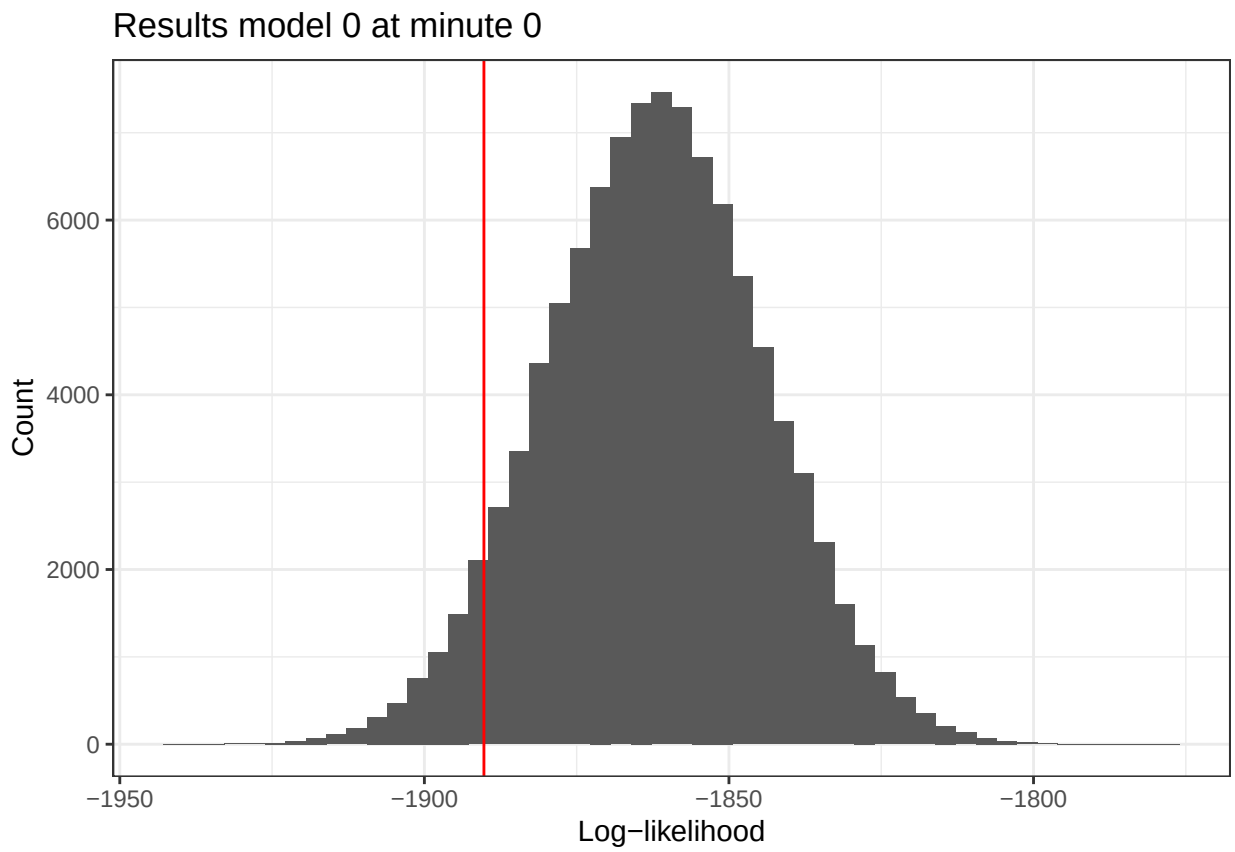
```
kable(tab_home_goals[1:4,], digits = 4)
```

	0	1	2	3	4	5+
Observed	0.2282	0.3617	0.2465	0.1173	0.0350	0.0113
Model 0 (min 0)	0.2571	0.3299	0.2302	0.1150	0.0459	0.0220
Model 3 (min 0)	0.2387	0.3426	0.2466	0.1169	0.0409	0.0143
Model 8 (min 0)	0.2348	0.3510	0.2462	0.1122	0.0396	0.0162

```
kable(tab_away_goals[1:4,], digits = 4)
```

	0	1	2	3	4	5+
Observed	0.3870	0.3671	0.1792	0.0474	0.0151	0.0043
Model 0 (min 0)	0.4056	0.3526	0.1663	0.0561	0.0151	0.0043
Model 3 (min 0)	0.3932	0.3677	0.1717	0.0529	0.0120	0.0025
Model 8 (min 0)	0.3923	0.3701	0.1712	0.0517	0.0119	0.0028

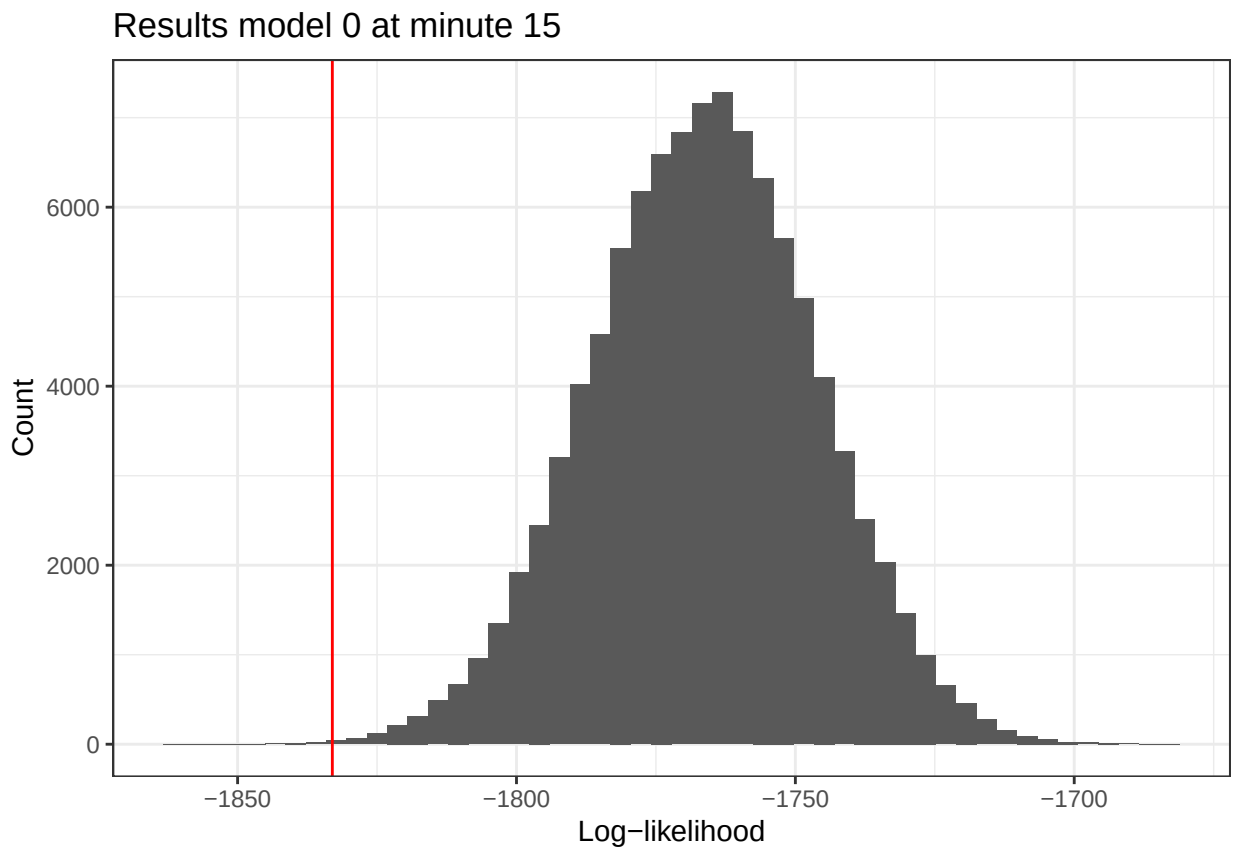
```
tibble(x = sims$pred_0$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 0")
```




```
sum(sims$pred_0$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_0) /
  length(sims$pred_0$loglik_results_mod_0)
```

```
## [1] 0.06041
```

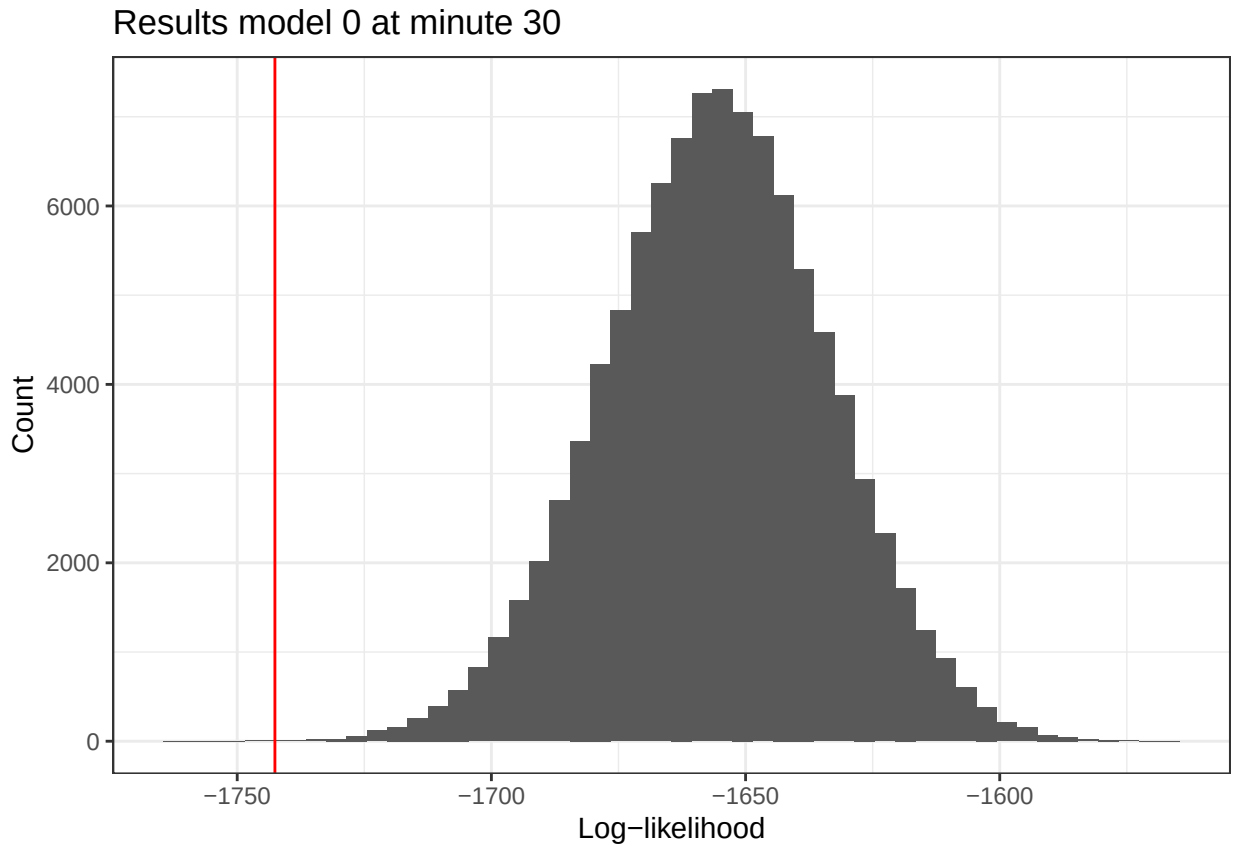
```
tibble(x = sims$pred_15$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 15")
```




```
sum(sims$pred_15$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_15) /
  length(sims$pred_15$loglik_results_mod_0)
```

```
## [1] 0.00064
```

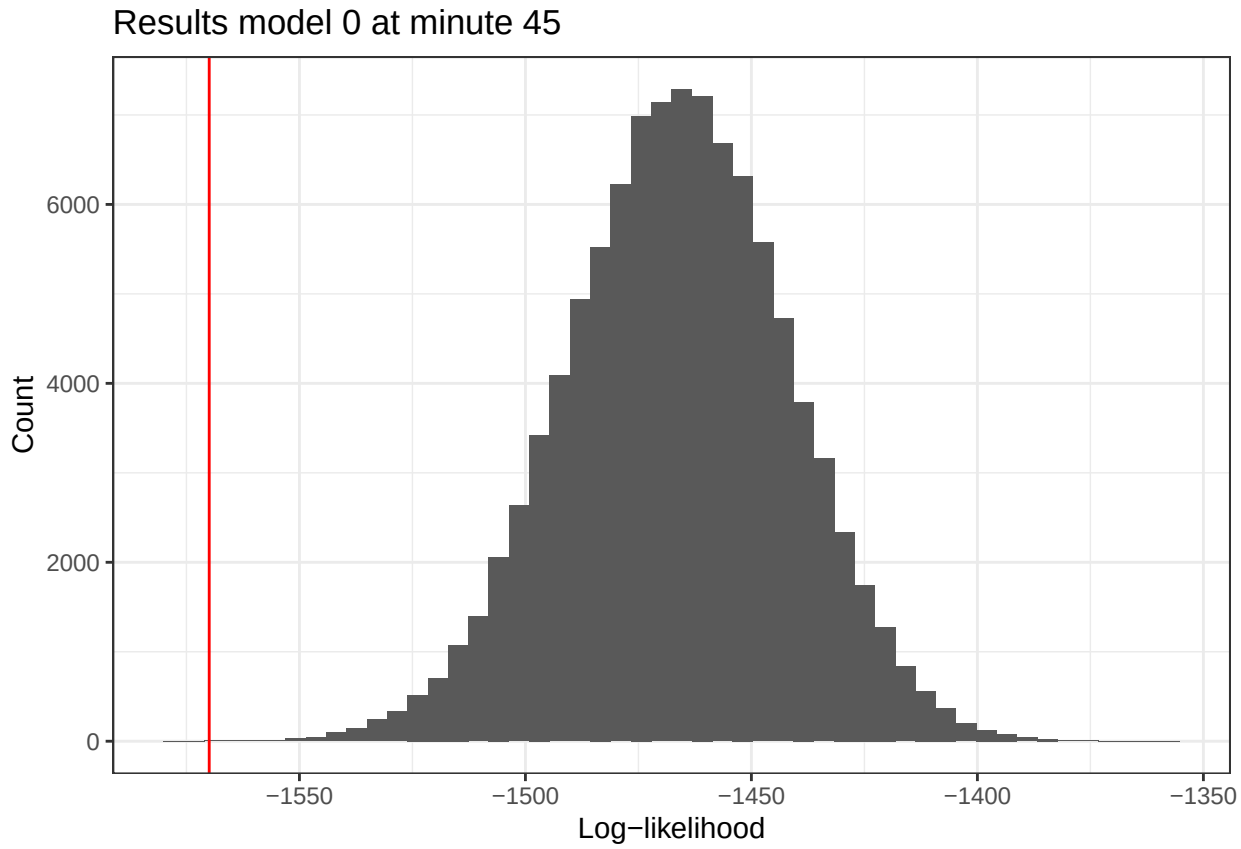
```
tibble(x = sims$pred_30$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 30")
```




```
sum(sims$pred_30$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_30) /
  length(sims$pred_30$loglik_results_mod_0)
```

```
## [1] 0.00011
```

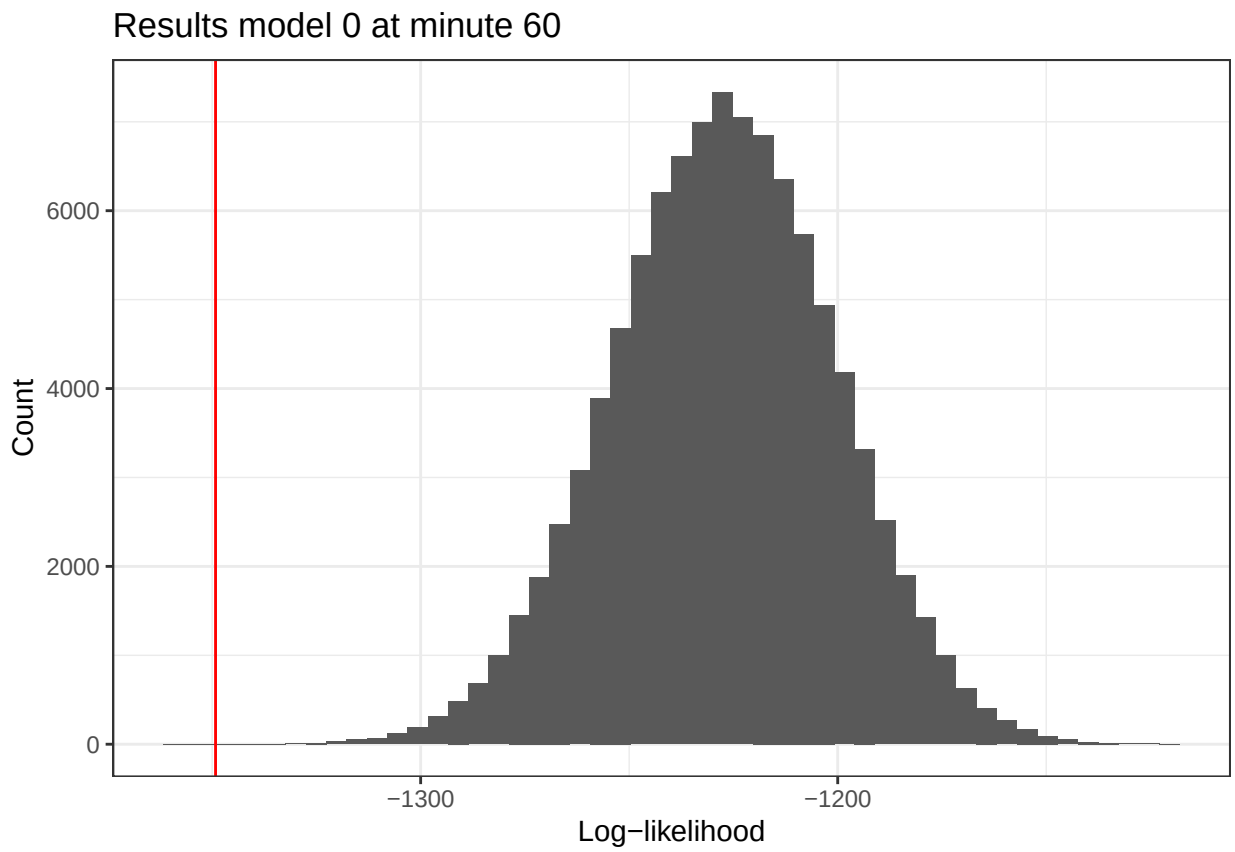
```
tibble(x = sims$pred_45$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 45")
```




```
sum(sims$pred_45$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_45) /
  length(sims$pred_45$loglik_results_mod_0)
```

```
## [1] 5e-05
```

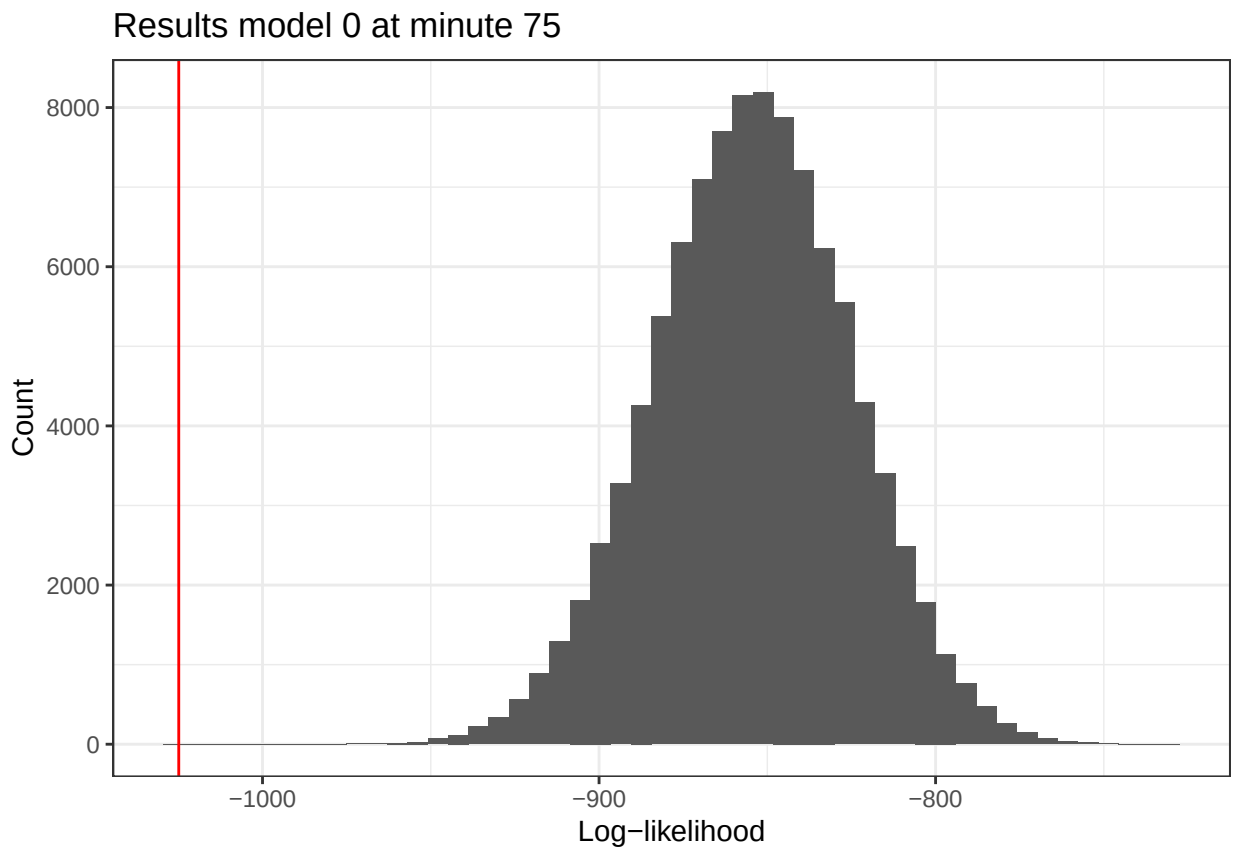
```
tibble(x = sims$pred_60$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 60")
```




```
sum(sims$pred_60$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_60) /
  length(sims$pred_60$loglik_results_mod_0)
```

```
## [1] 2e-05
```

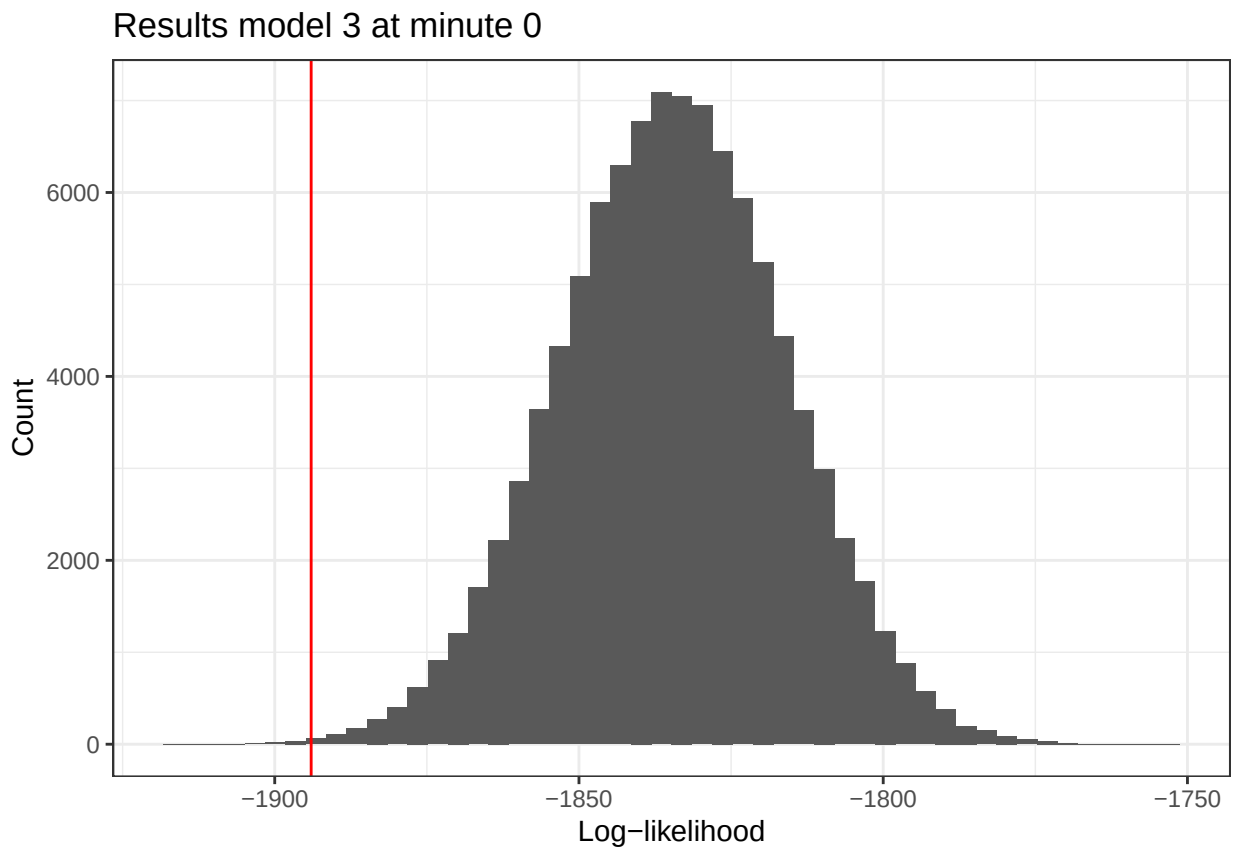
```
tibble(x = sims$pred_75$loglik_results_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_0_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 0 at minute 75")
```




```
sum(sims$pred_75$loglik_results_mod_0 <= loglik_observed_results_mod_0_pred_75) /
  length(sims$pred_75$loglik_results_mod_0)
```

```
## [1] 0
```

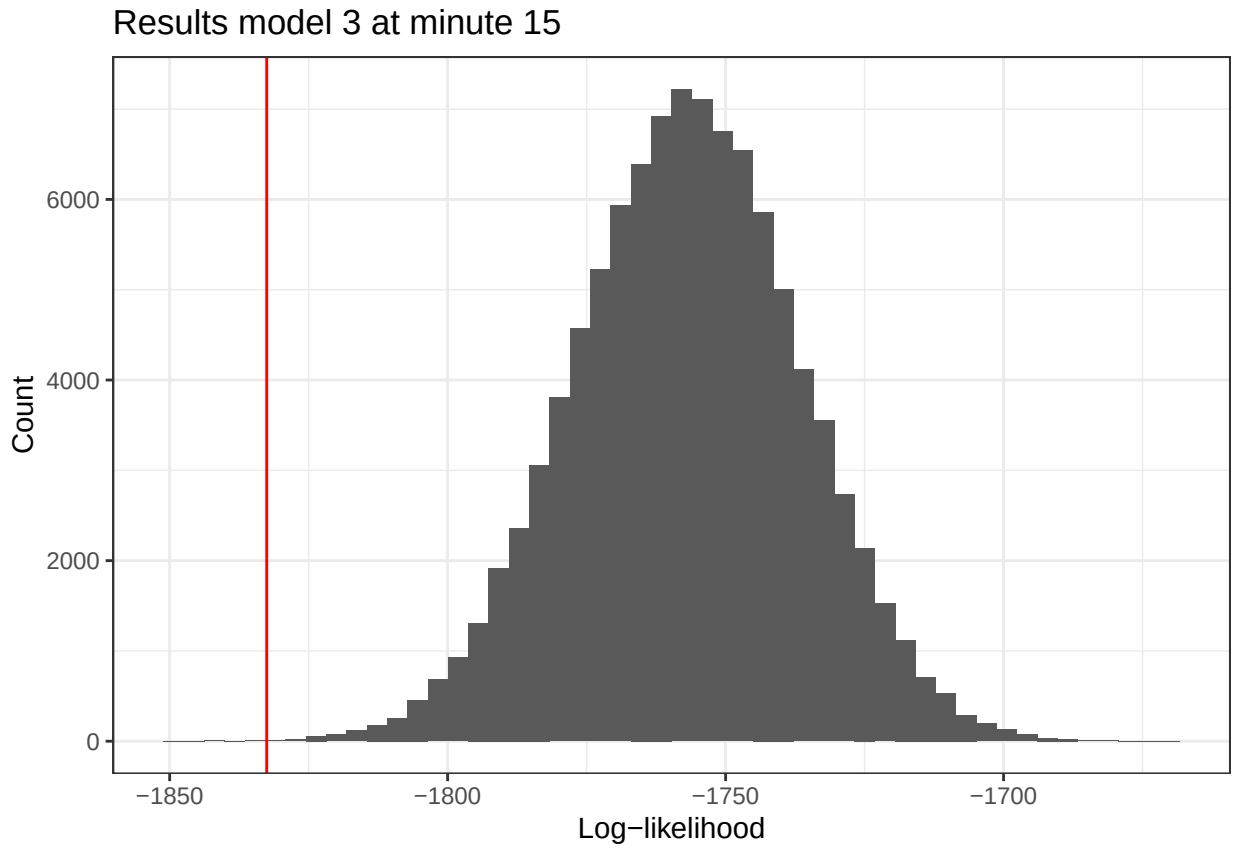
```
tibble(x = sims$pred_0$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 0")
```




```
sum(sims$pred_0$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_0) /
  length(sims$pred_0$loglik_results_mod_3)
```

```
## [1] 0.00078
```

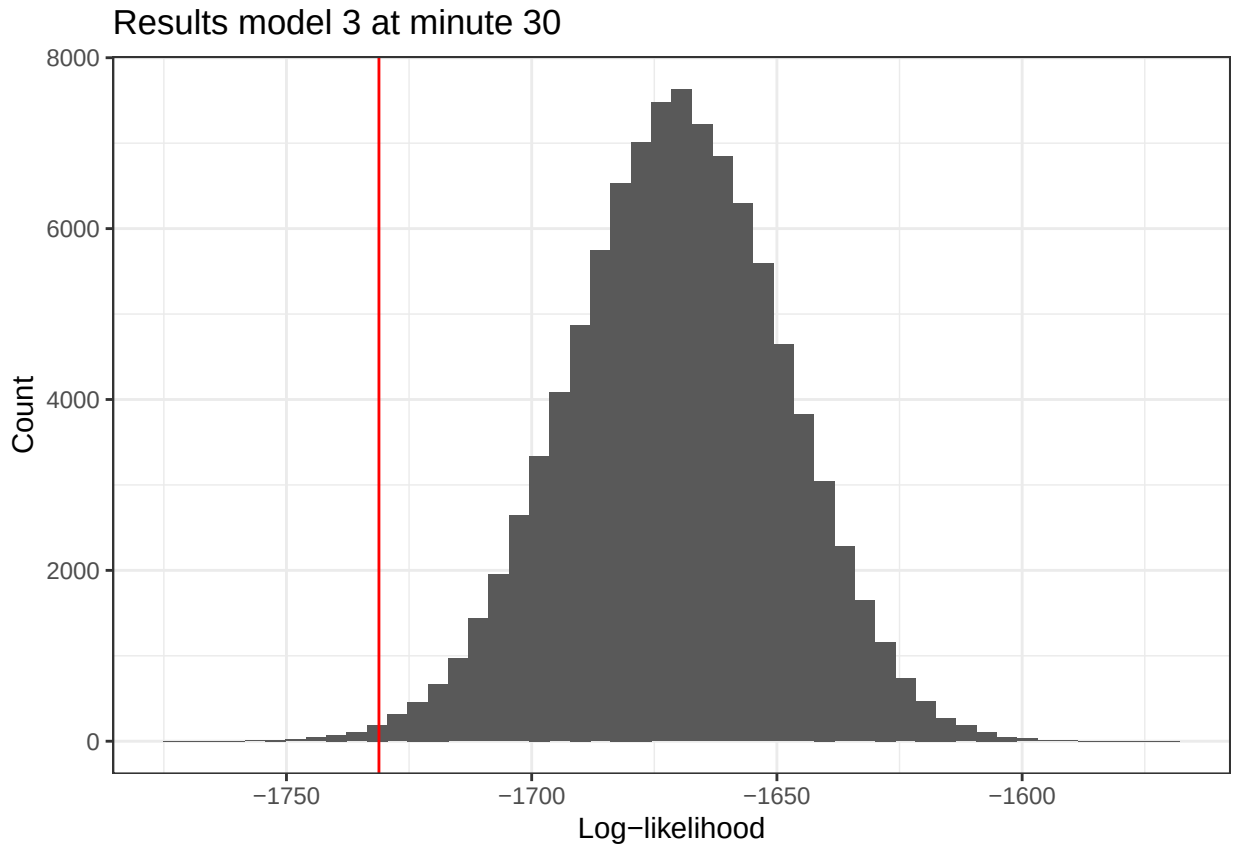
```
tibble(x = sims$pred_15$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 15")
```




```
sum(sims$pred_15$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_15) /
  length(sims$pred_15$loglik_results_mod_3)
```

```
## [1] 0.00016
```

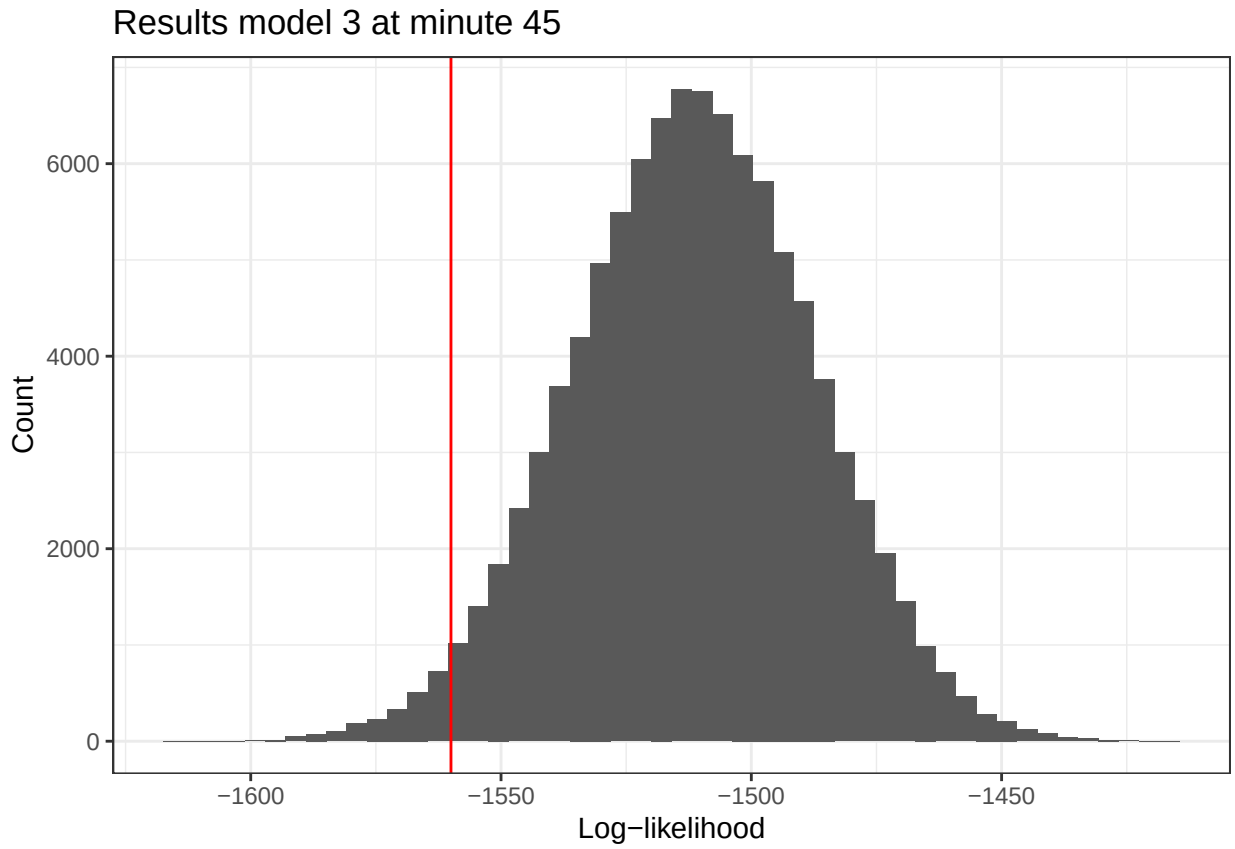
```
tibble(x = sims$pred_30$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 30")
```




```
sum(sims$pred_30$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_30) /
  length(sims$pred_30$loglik_results_mod_3)
```

```
## [1] 0.00368
```

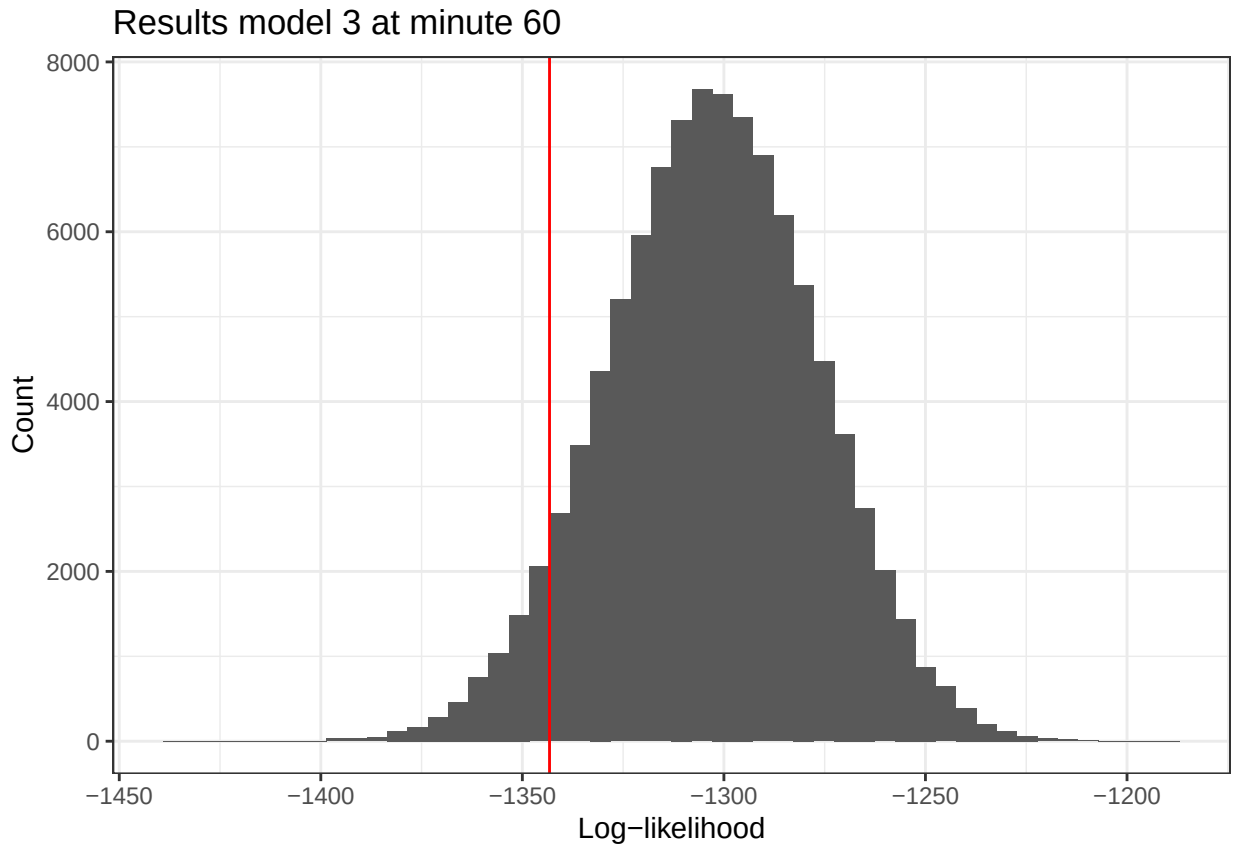
```
tibble(x = sims$pred_45$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 45")
```




```
sum(sims$pred_45$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_45) /
  length(sims$pred_45$loglik_results_mod_3)
```

```
## [1] 0.02366
```

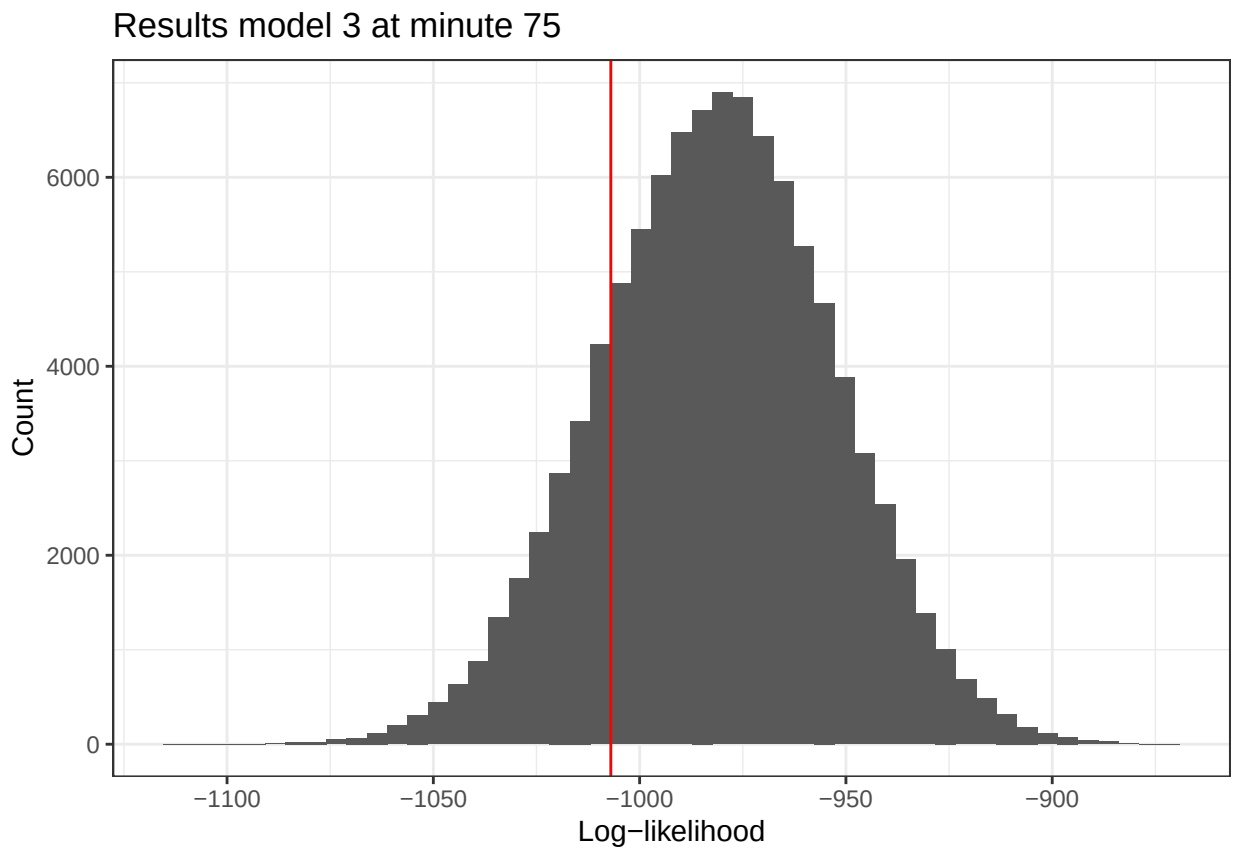
```
tibble(x = sims$pred_60$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 60")
```




```
sum(sims$pred_60$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_60) /
  length(sims$pred_60$loglik_results_mod_3)
```

```
## [1] 0.06465
```

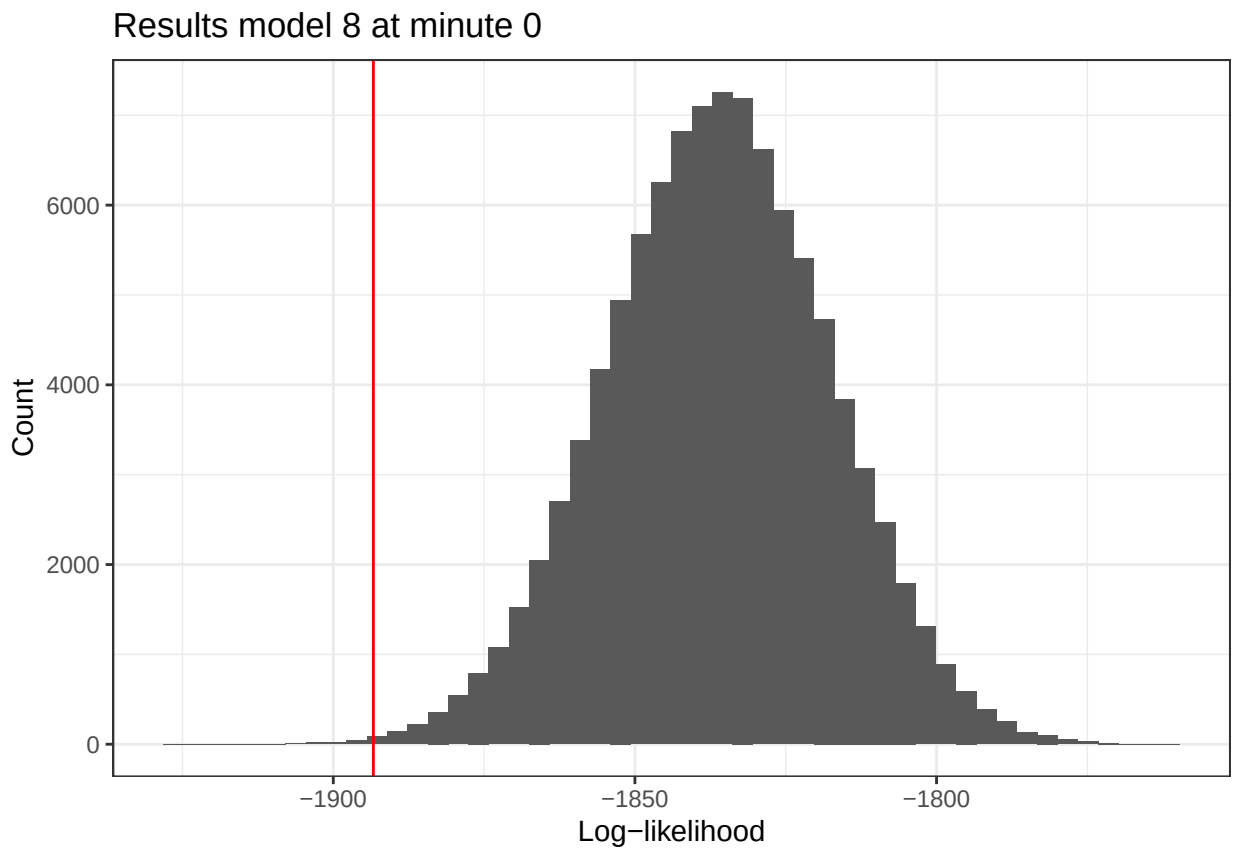
```
tibble(x = sims$pred_75$loglik_results_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_3_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 3 at minute 75")
```




```
sum(sims$pred_75$loglik_results_mod_3 <= loglik_observed_results_mod_3_pred_75) /
  length(sims$pred_75$loglik_results_mod_3)
```

```
## [1] 0.18622
```

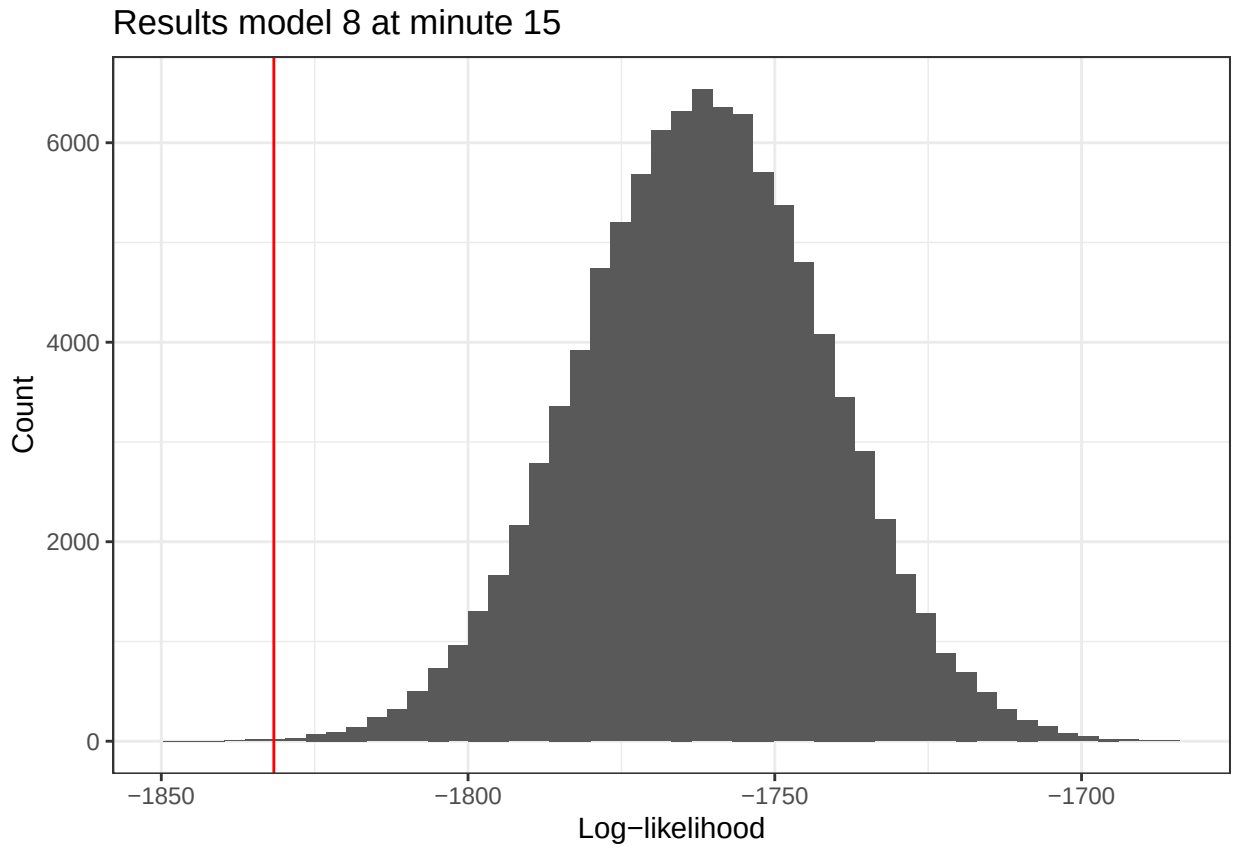
```
tibble(x = sims$pred_0$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 0")
```




```
sum(sims$pred_0$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_0) /
  length(sims$pred_0$loglik_results_mod_8)
```

```
## [1] 0.00119
```

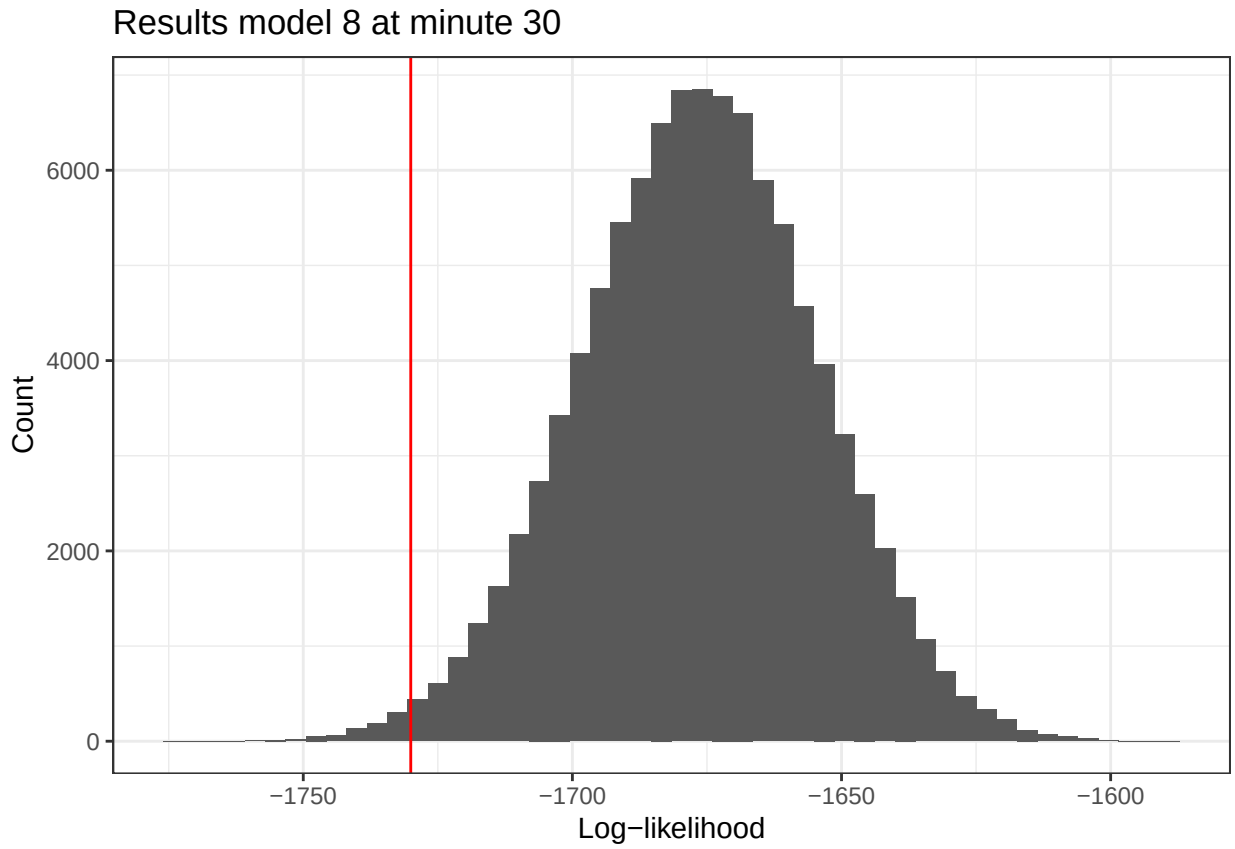
```
tibble(x = sims$pred_15$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 15")
```




```
sum(sims$pred_15$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_15) /
  length(sims$pred_15$loglik_results_mod_8)
```

```
## [1] 0.00036
```

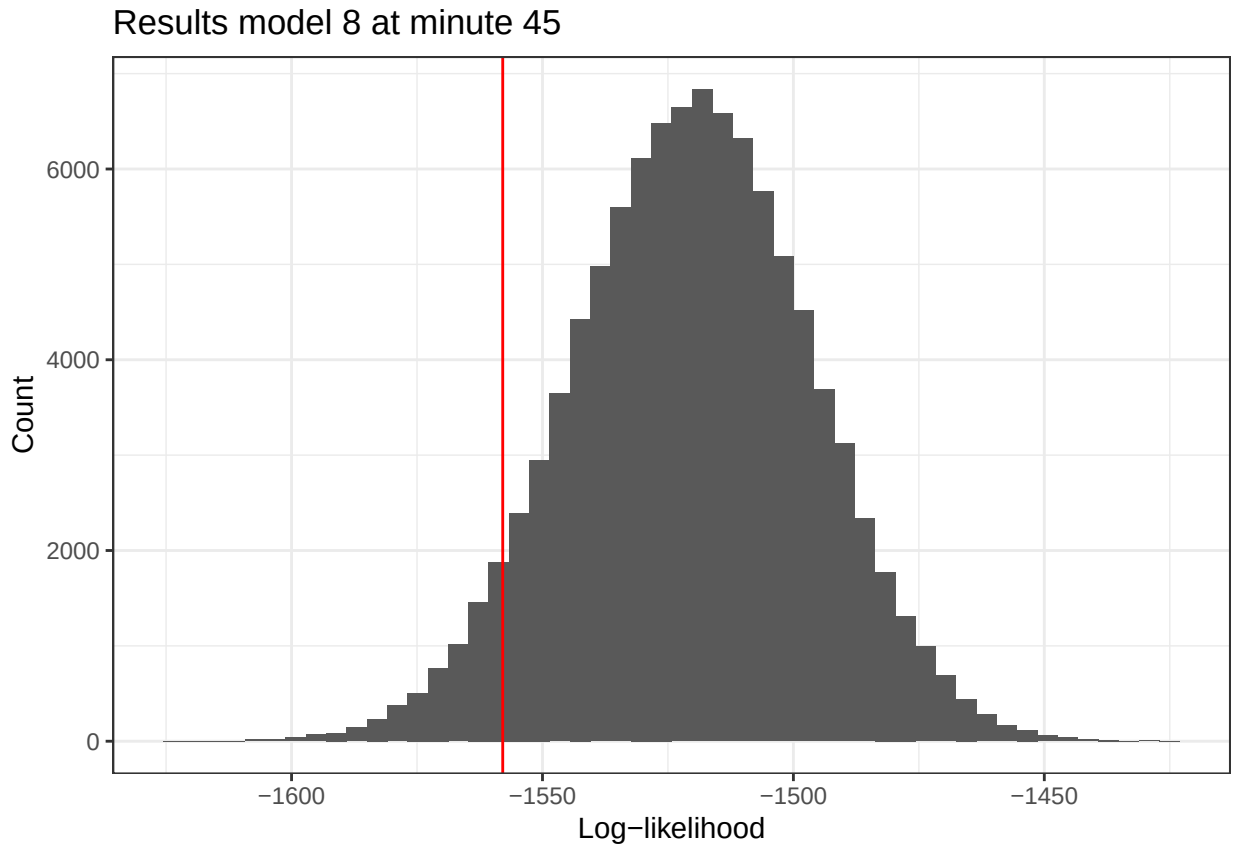
```
tibble(x = sims$pred_30$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 30")
```




```
sum(sims$pred_30$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_30) /
  length(sims$pred_30$loglik_results_mod_8)
```

```
## [1] 0.00847
```

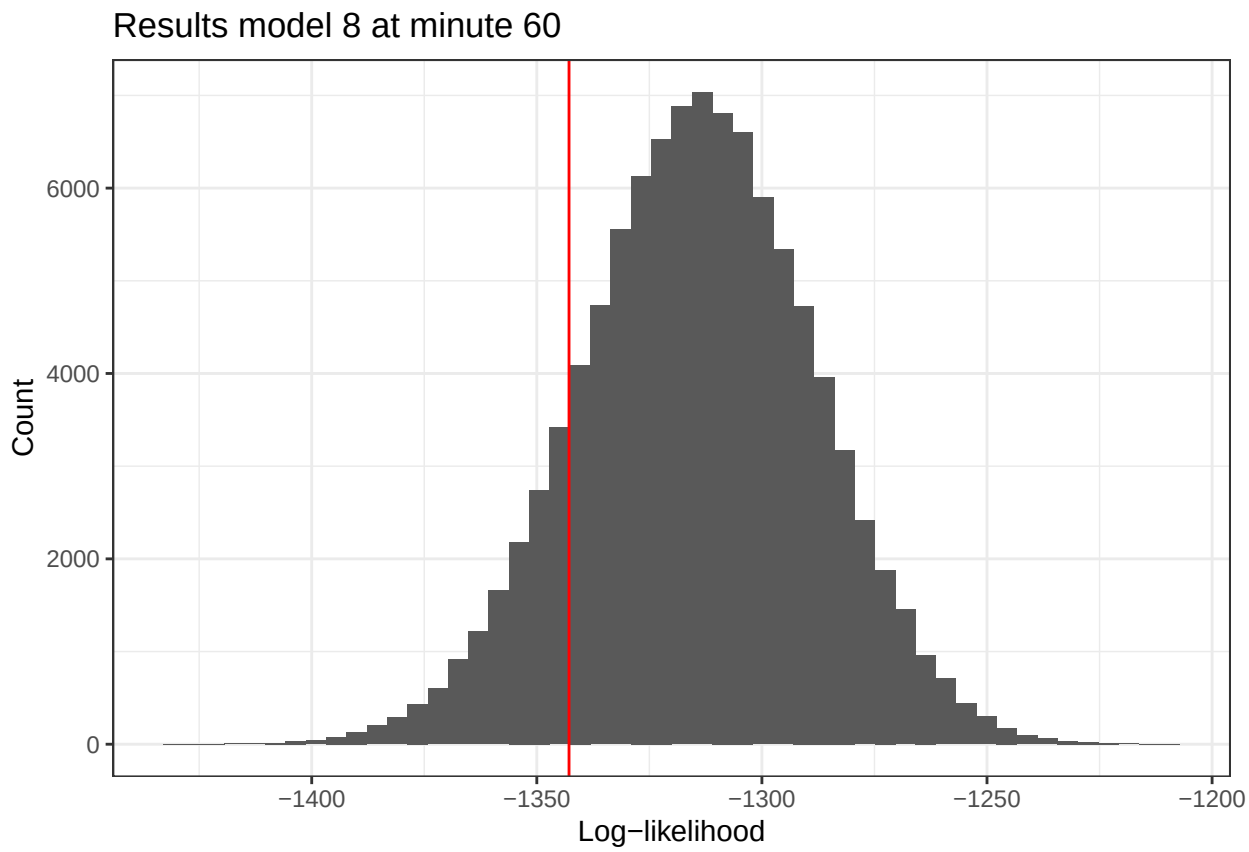
```
tibble(x = sims$pred_45$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 45")
```




```
sum(sims$pred_45$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_45) /
  length(sims$pred_45$loglik_results_mod_8)
```

```
## [1] 0.05918
```

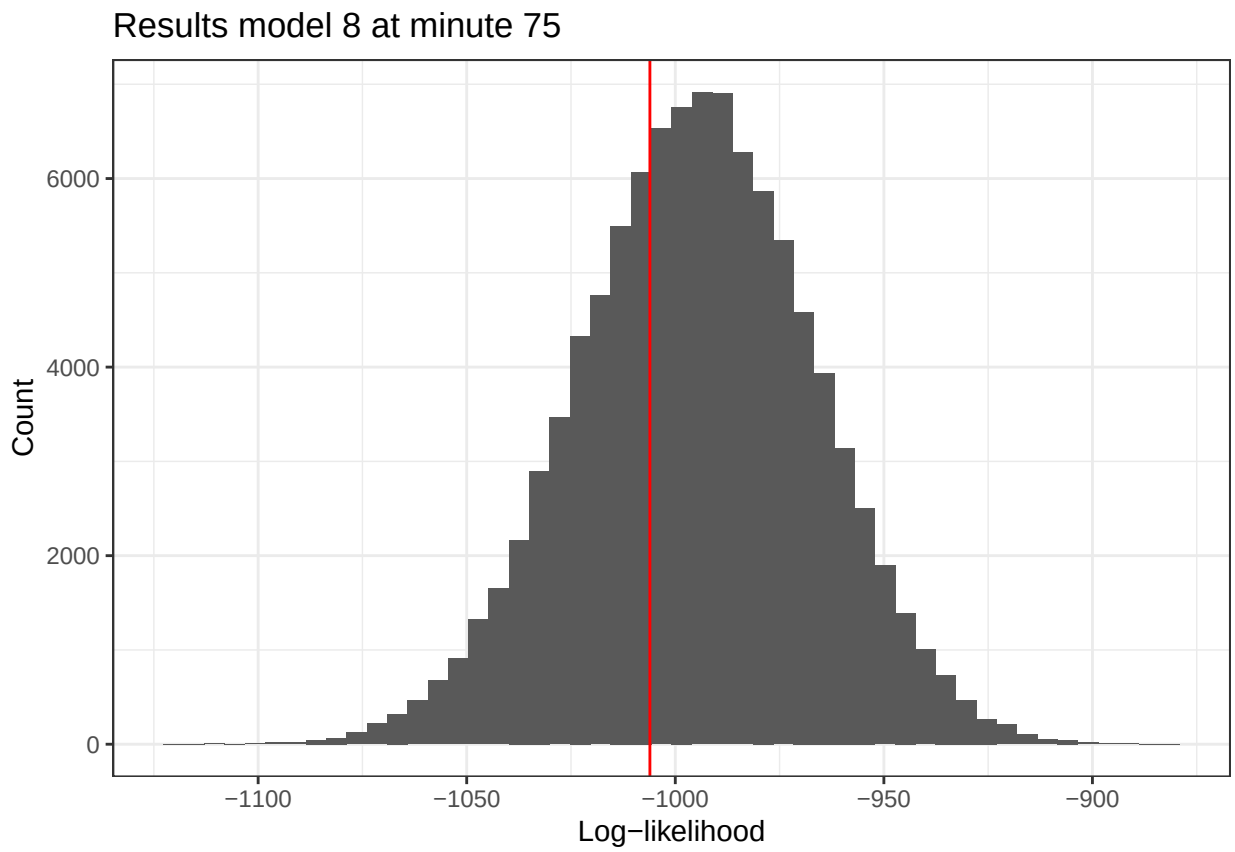
```
tibble(x = sims$pred_60$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 60")
```




```
sum(sims$pred_60$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_60) /
  length(sims$pred_60$loglik_results_mod_8)
```

```
## [1] 0.13774
```

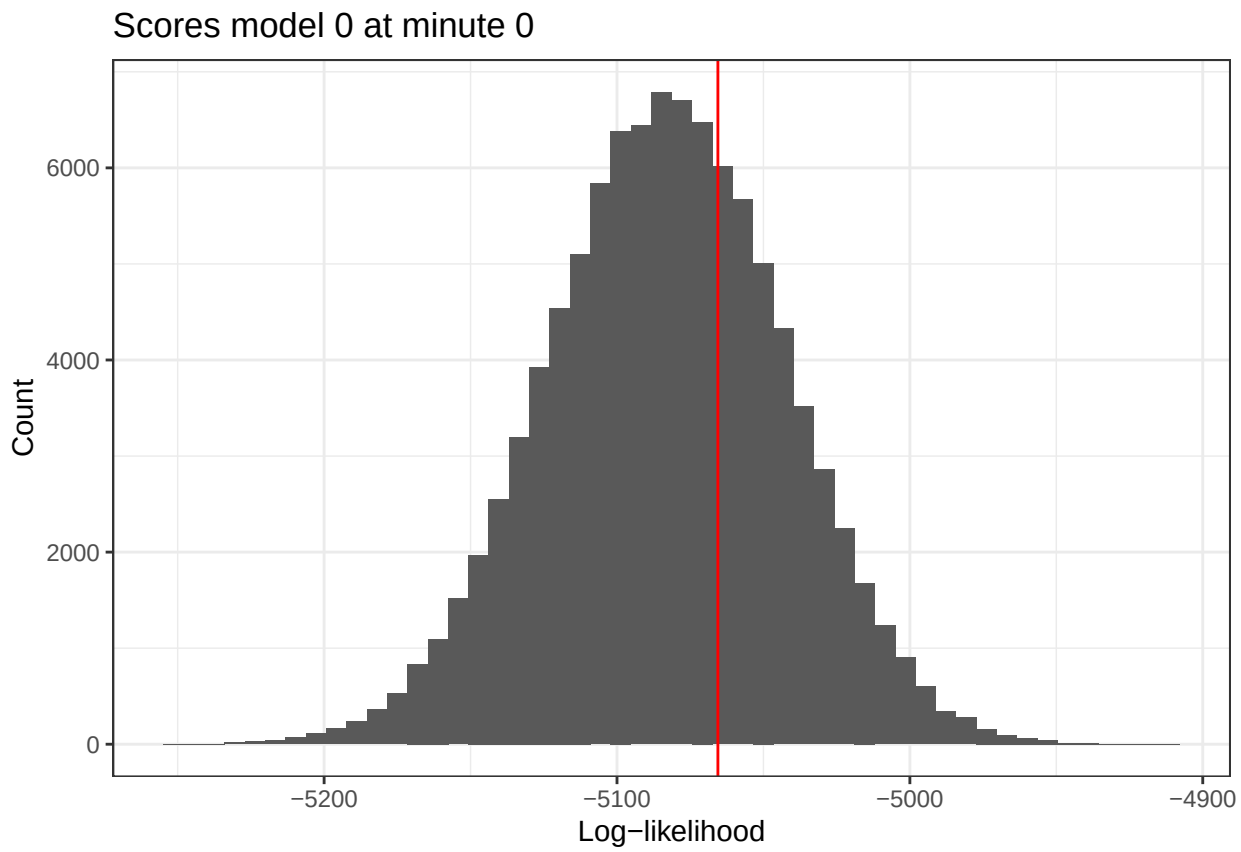
```
tibble(x = sims$pred_75$loglik_results_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_results_mod_8_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Results model 8 at minute 75")
```




```
sum(sims$pred_75$loglik_results_mod_8 <= loglik_observed_results_mod_8_pred_75) /
  length(sims$pred_75$loglik_results_mod_8)
```

```
## [1] 0.34585
```

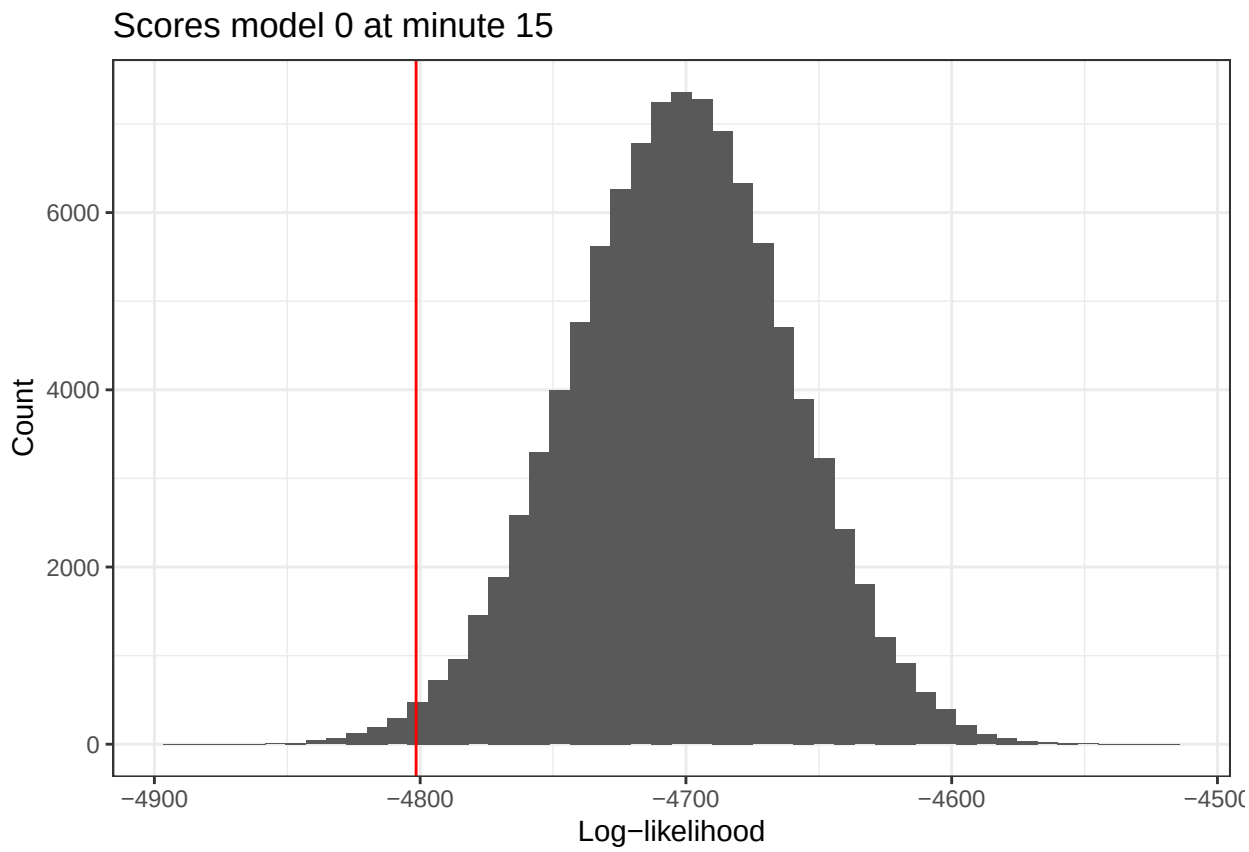
```
tibble(x = sims$pred_0$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 0")
```




```
sum(sims$pred_0$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_0) /
  length(sims$pred_0$loglik_scores_mod_0)
```

```
## [1] 0.66653
```

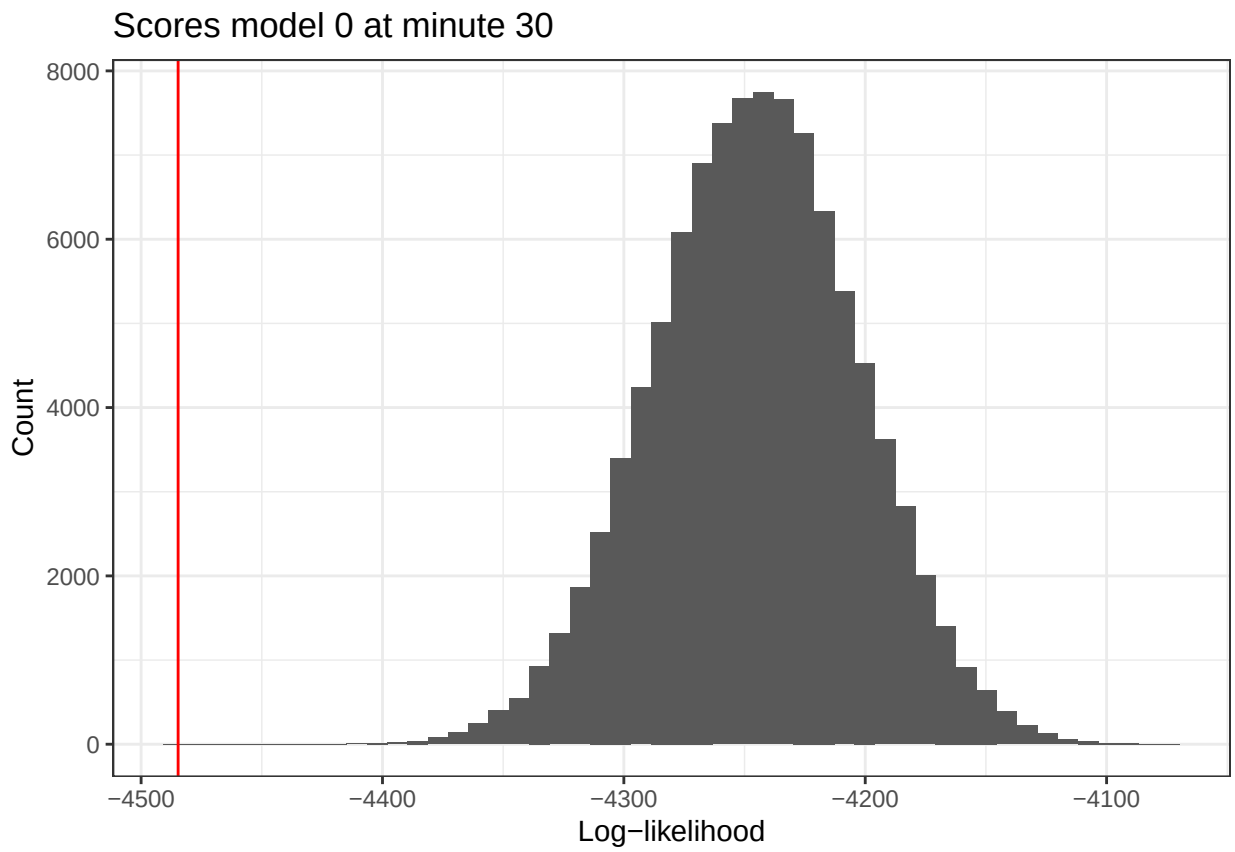
```
tibble(x = sims$pred_15$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 15")
```




```
sum(sims$pred_15$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_15) /
  length(sims$pred_15$loglik_scores_mod_0)
```

```
## [1] 0.00952
```

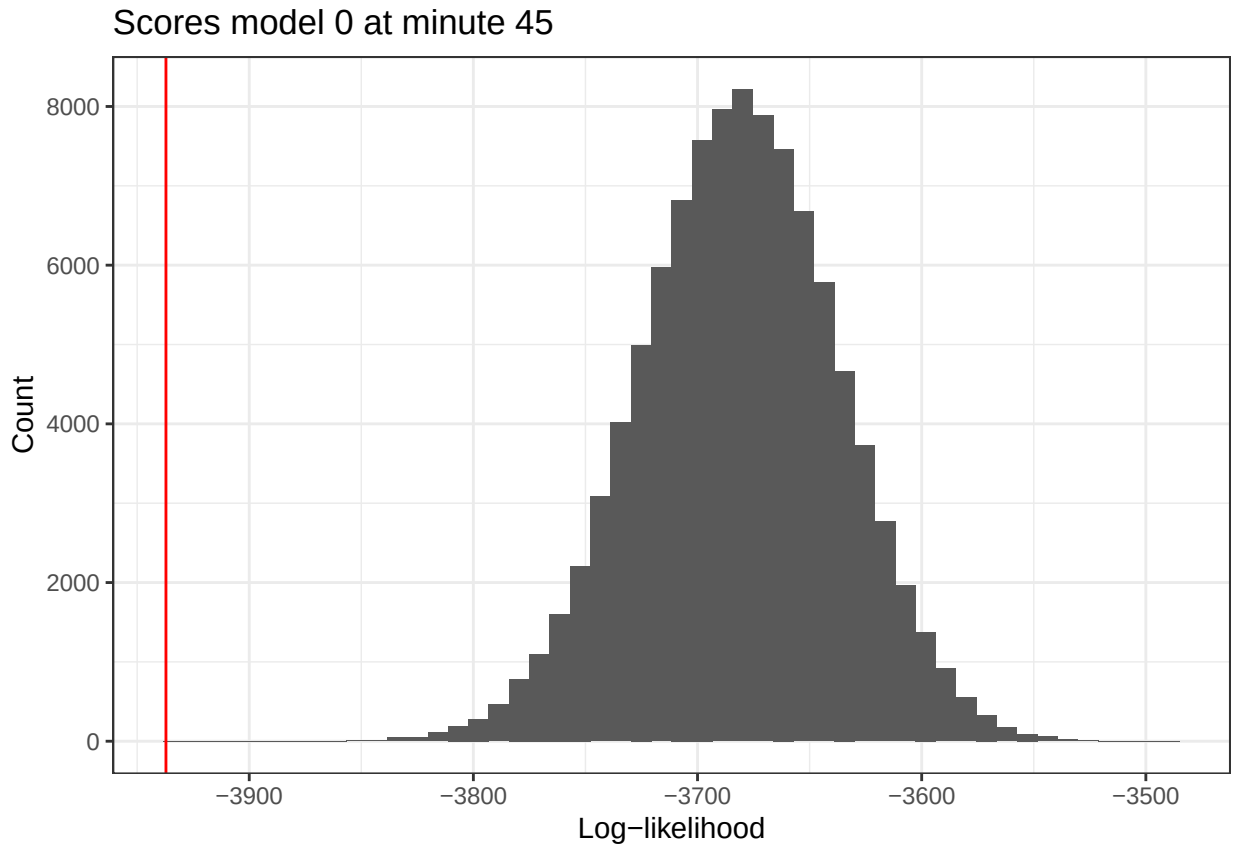
```
tibble(x = sims$pred_30$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 30")
```




```
sum(sims$pred_30$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_30) /
  length(sims$pred_30$loglik_scores_mod_0)
```

```
## [1] 0
```

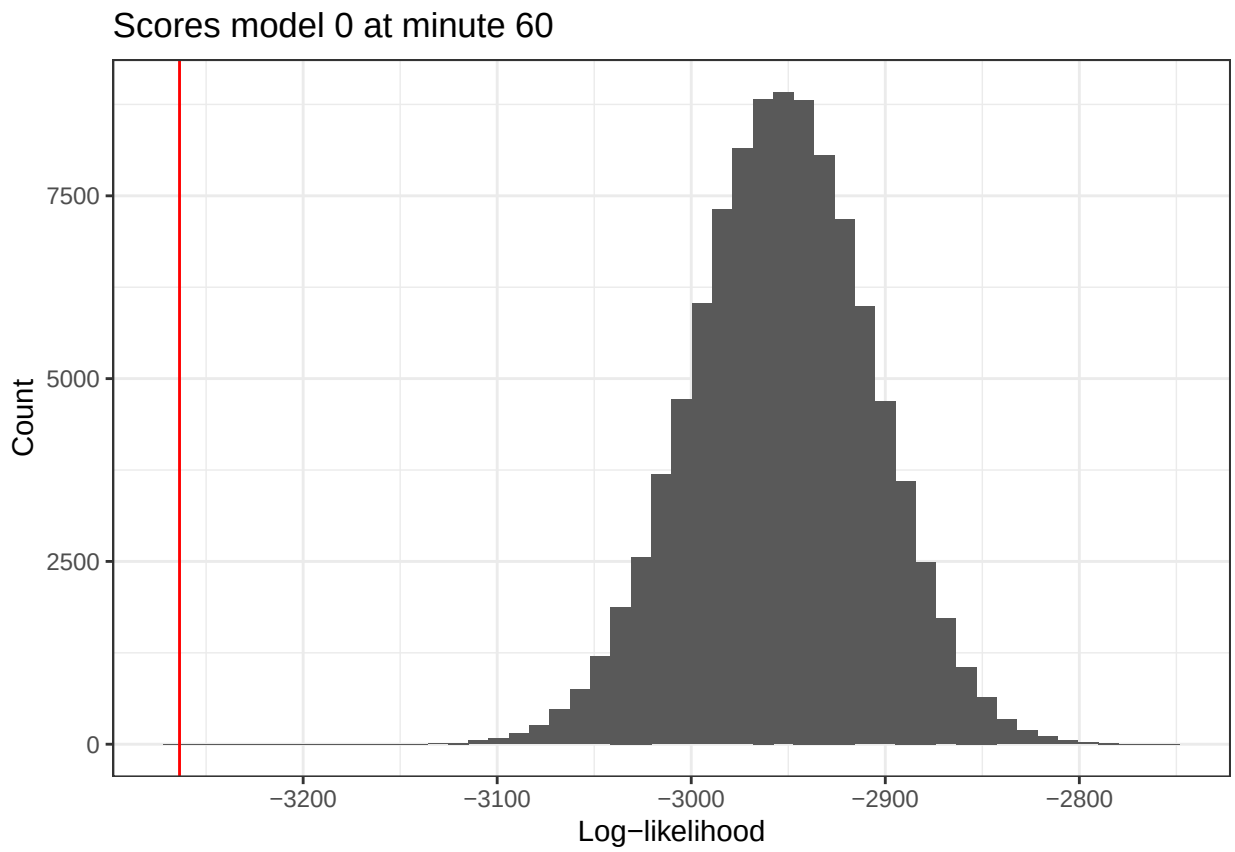
```
tibble(x = sims$pred_45$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 45")
```




```
sum(sims$pred_45$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_45) /
  length(sims$pred_45$loglik_scores_mod_0)
```

```
## [1] 0
```

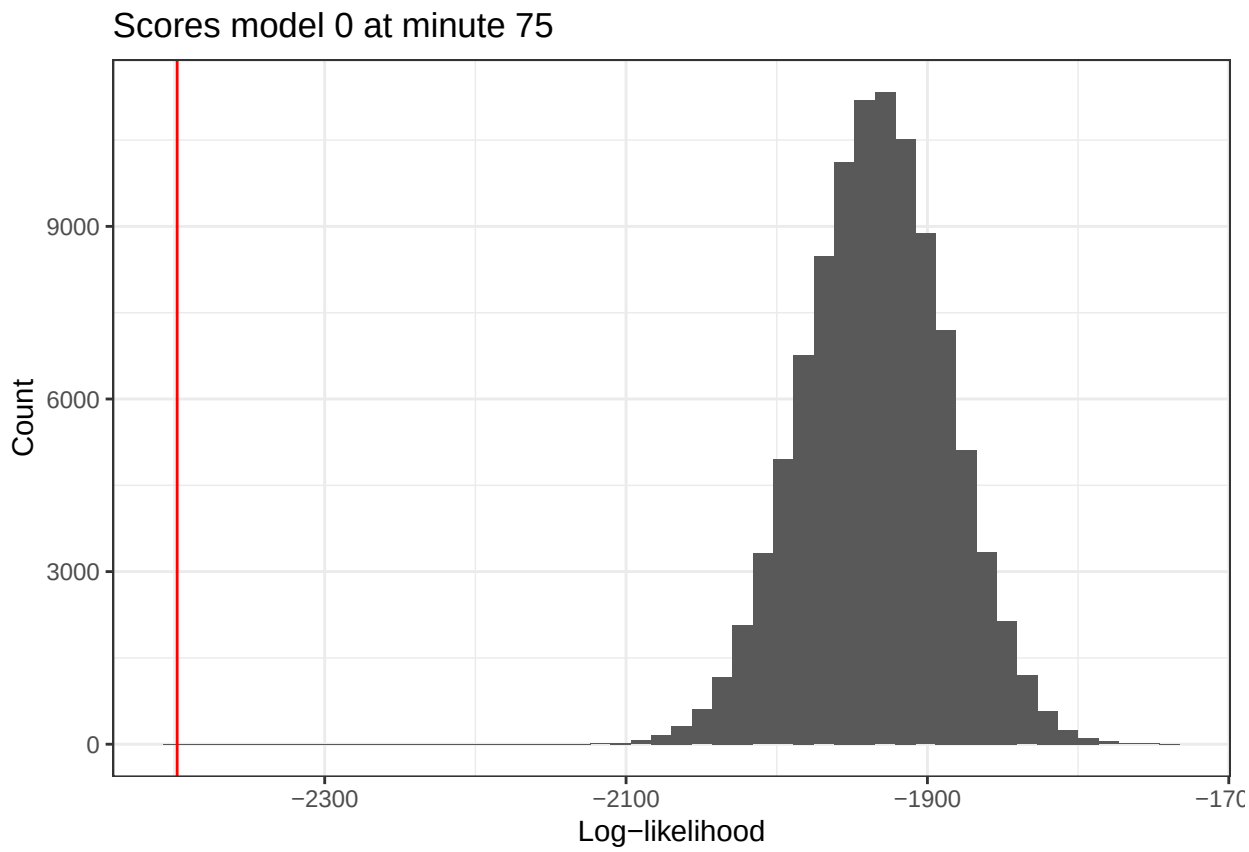
```
tibble(x = sims$pred_60$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 60")
```




```
sum(sims$pred_60$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_60) /
  length(sims$pred_60$loglik_scores_mod_0)
```

```
## [1] 0
```

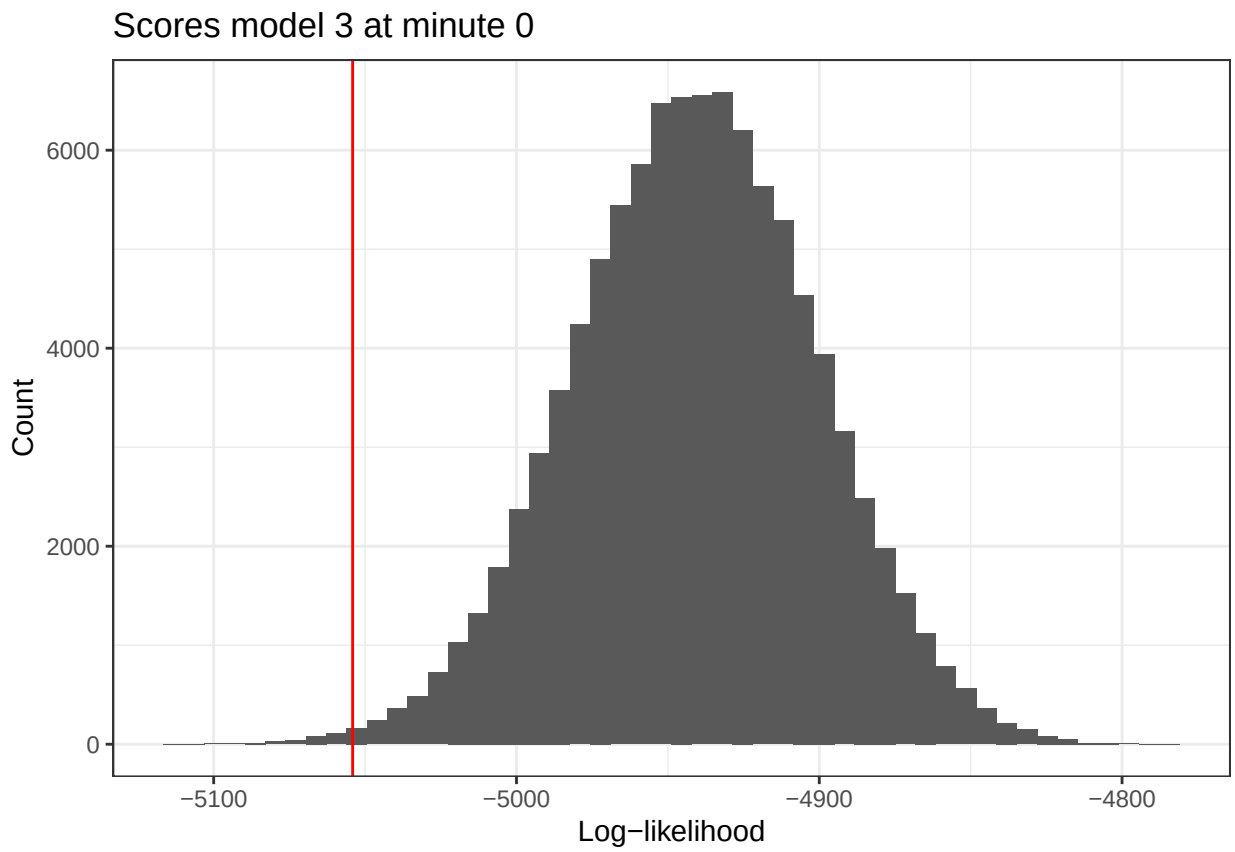
```
tibble(x = sims$pred_75$loglik_scores_mod_0) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_0_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 0 at minute 75")
```




```
sum(sims$pred_75$loglik_scores_mod_0 <= loglik_observed_scores_mod_0_pred_75) /
  length(sims$pred_75$loglik_scores_mod_0)
```

```
## [1] 0
```

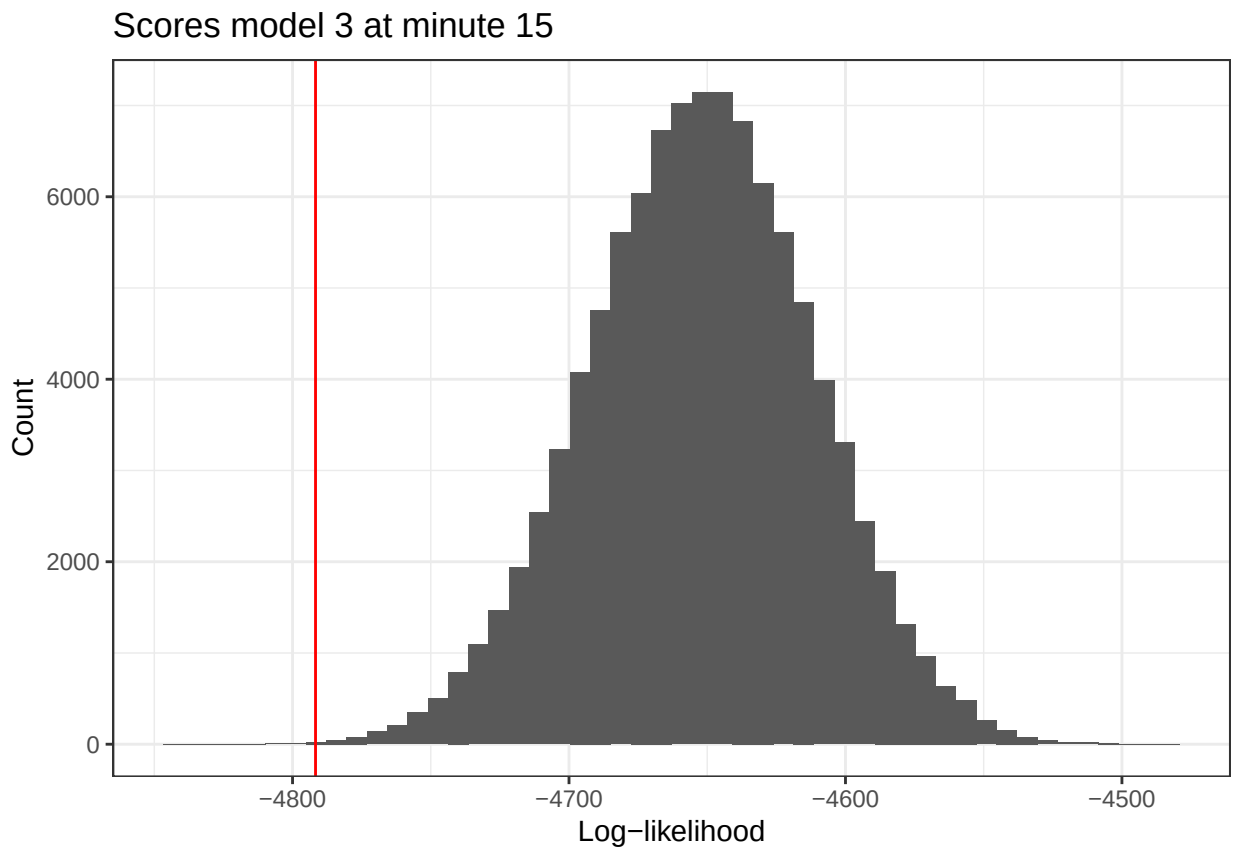
```
tibble(x = sims$pred_0$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 0")
```




```
sum(sims$pred_0$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_0) /
  length(sims$pred_0$loglik_scores_mod_3)
```

```
## [1] 0.00326
```

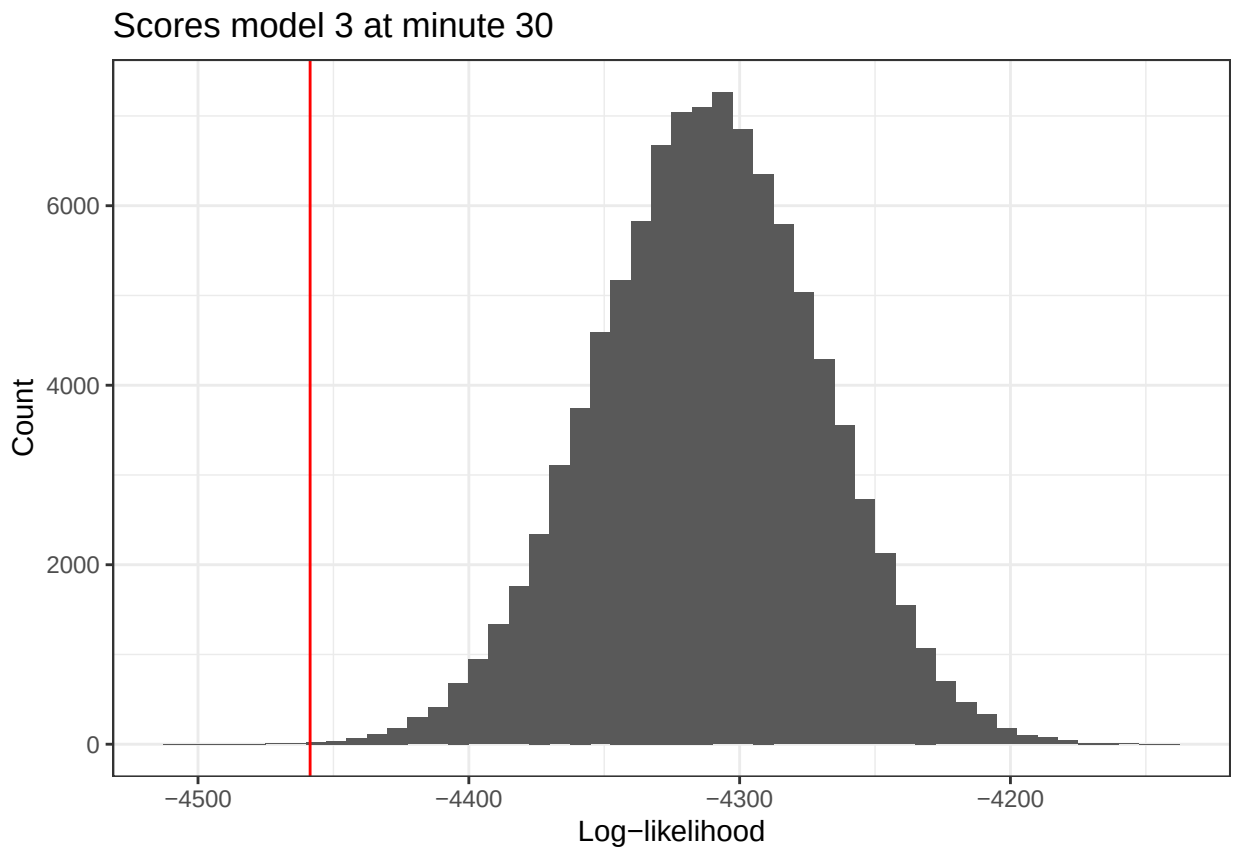
```
tibble(x = sims$pred_15$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 15")
```




```
sum(sims$pred_15$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_15) /
  length(sims$pred_15$loglik_scores_mod_3)
```

```
## [1] 0.00031
```

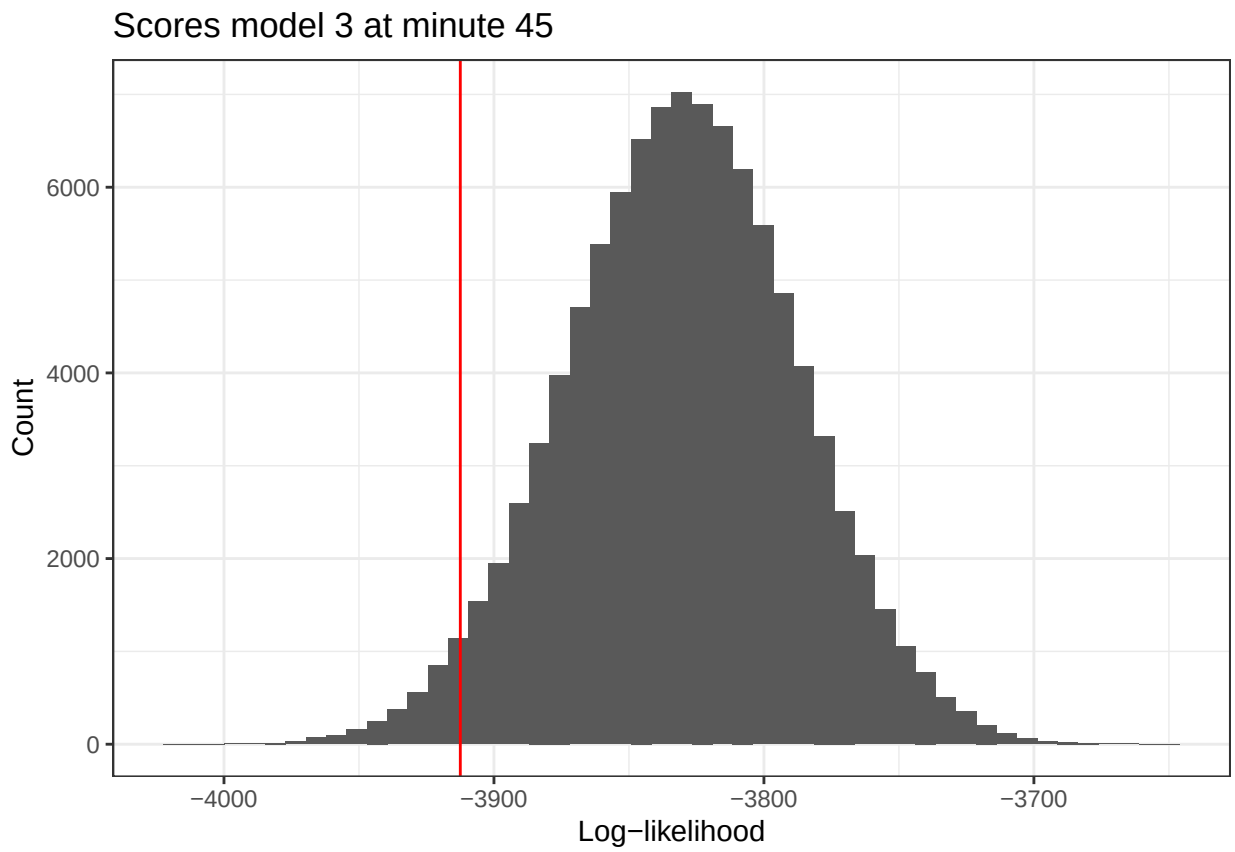
```
tibble(x = sims$pred_30$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 30")
```




```
sum(sims$pred_30$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_30) /
  length(sims$pred_30$loglik_scores_mod_3)
```

```
## [1] 0.00024
```

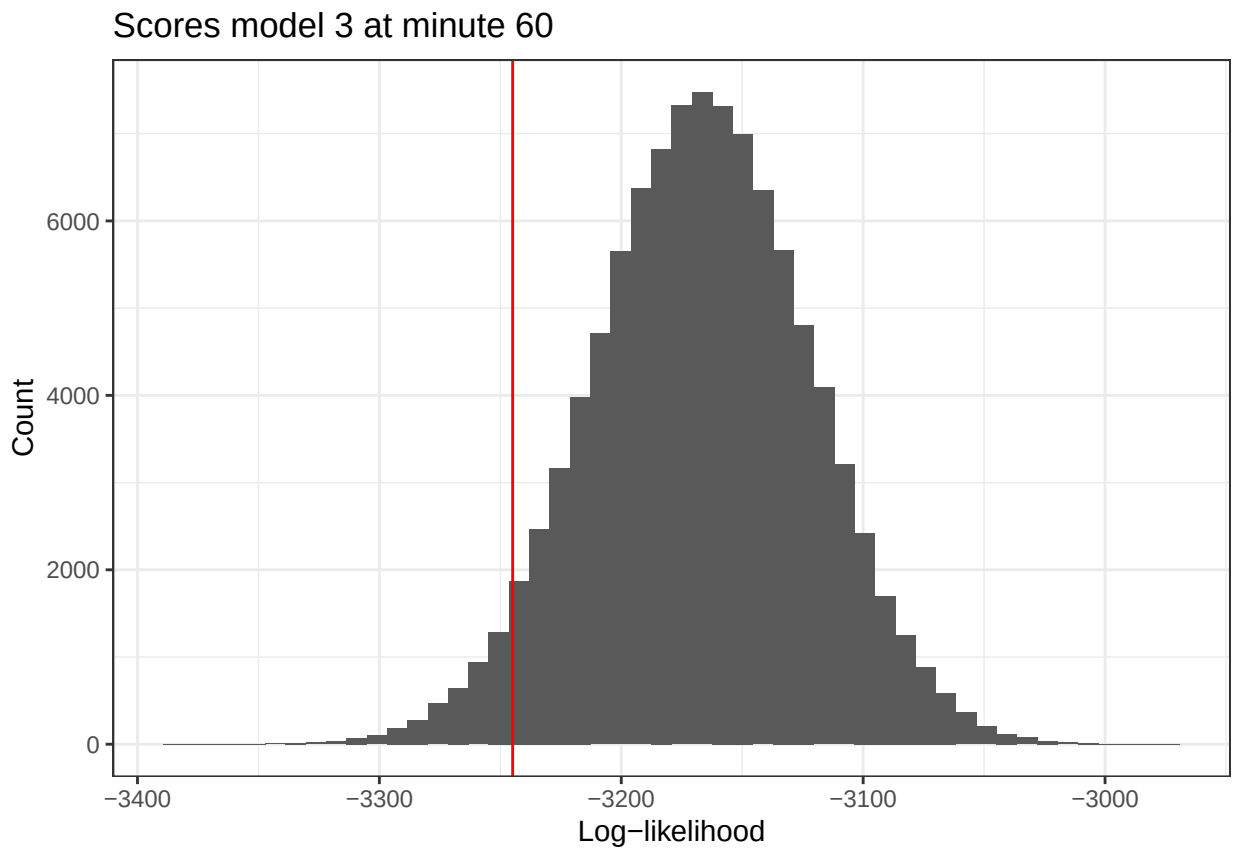
```
tibble(x = sims$pred_45$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 45")
```




```
sum(sims$pred_45$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_45) /
  length(sims$pred_45$loglik_scores_mod_3)
```

```
## [1] 0.03056
```

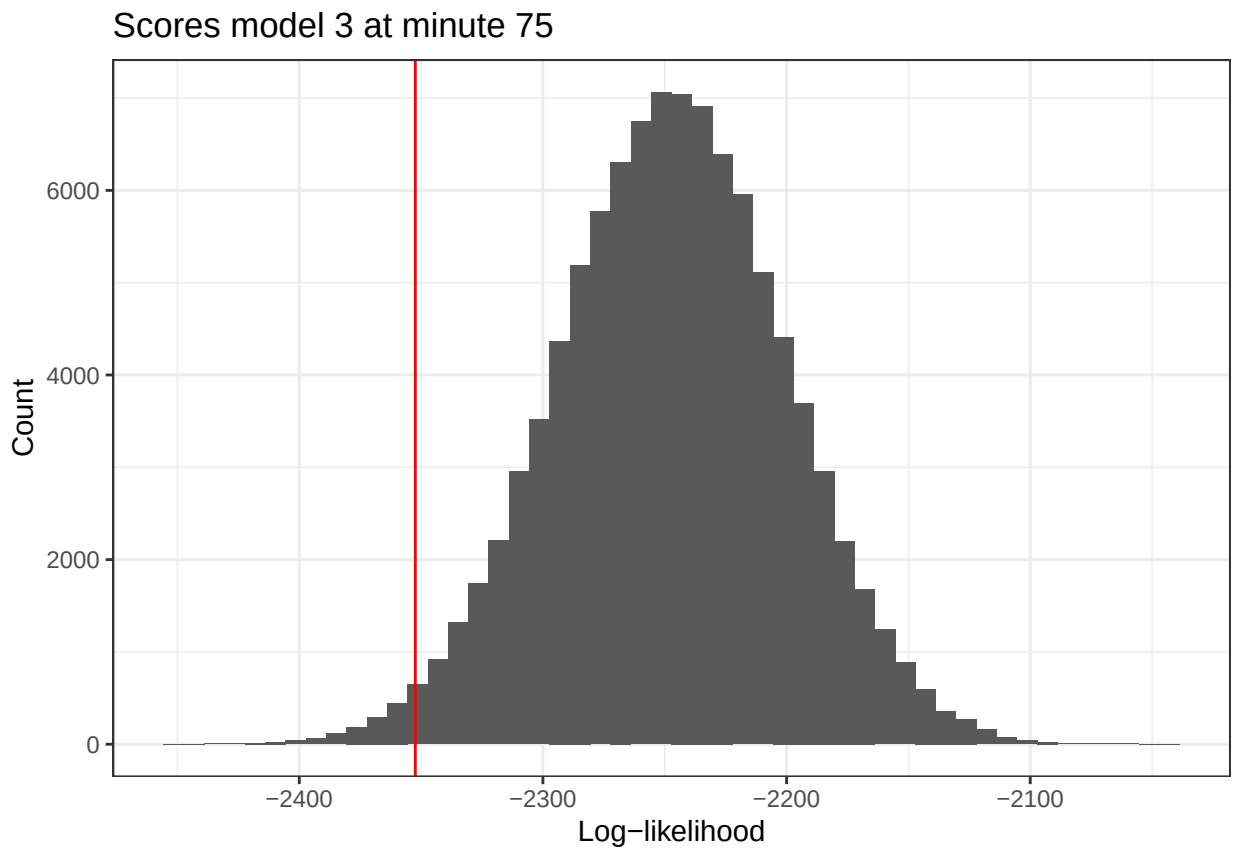
```
tibble(x = sims$pred_60$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 60")
```




```
sum(sims$pred_60$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_60) /
  length(sims$pred_60$loglik_scores_mod_3)
```

```
## [1] 0.04336
```

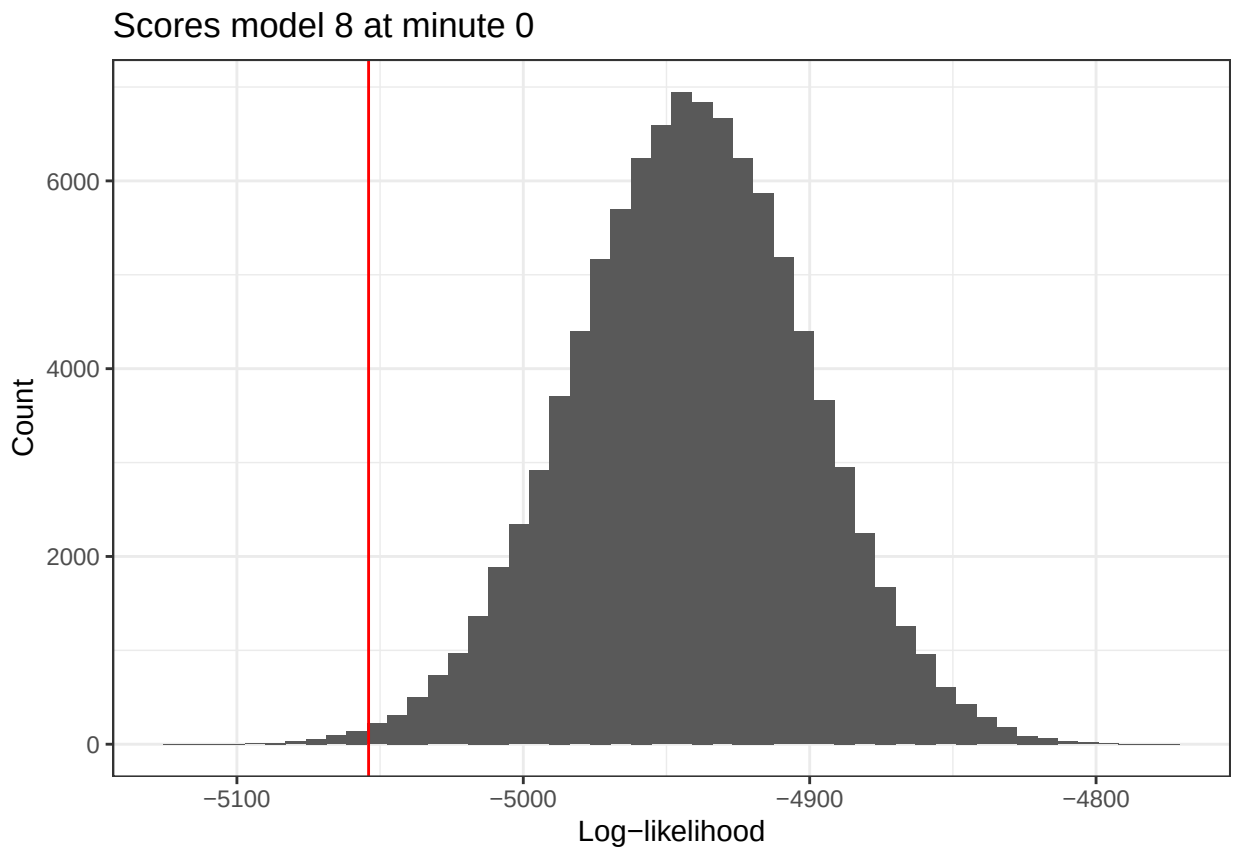
```
tibble(x = sims$pred_75$loglik_scores_mod_3) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_3_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 3 at minute 75")
```




```
sum(sims$pred_75$loglik_scores_mod_3 <= loglik_observed_scores_mod_3_pred_75) /
  length(sims$pred_75$loglik_scores_mod_3)
```

```
## [1] 0.01416
```

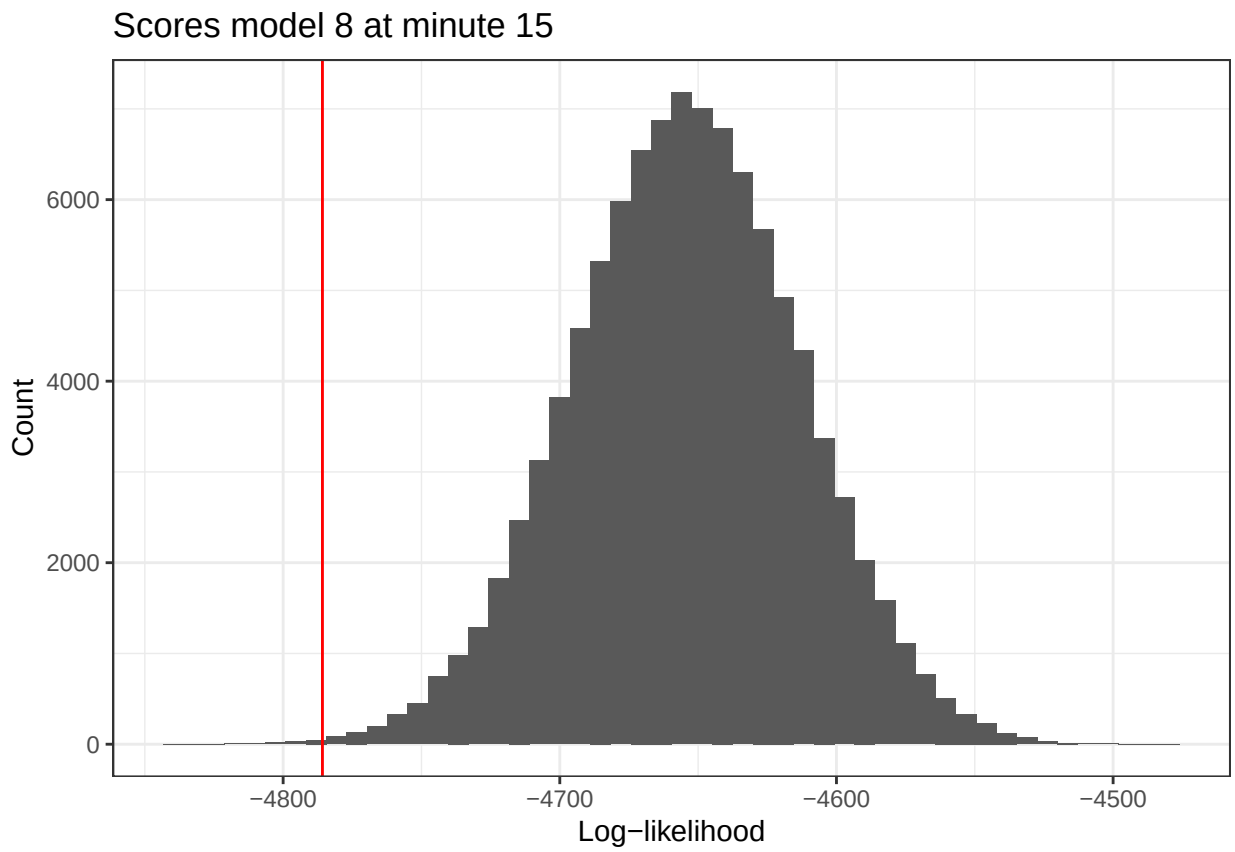
```
tibble(x = sims$pred_0$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_0, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 0")
```




```
sum(sims$pred_0$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_0) /
  length(sims$pred_0$loglik_scores_mod_8)
```

```
## [1] 0.00367
```

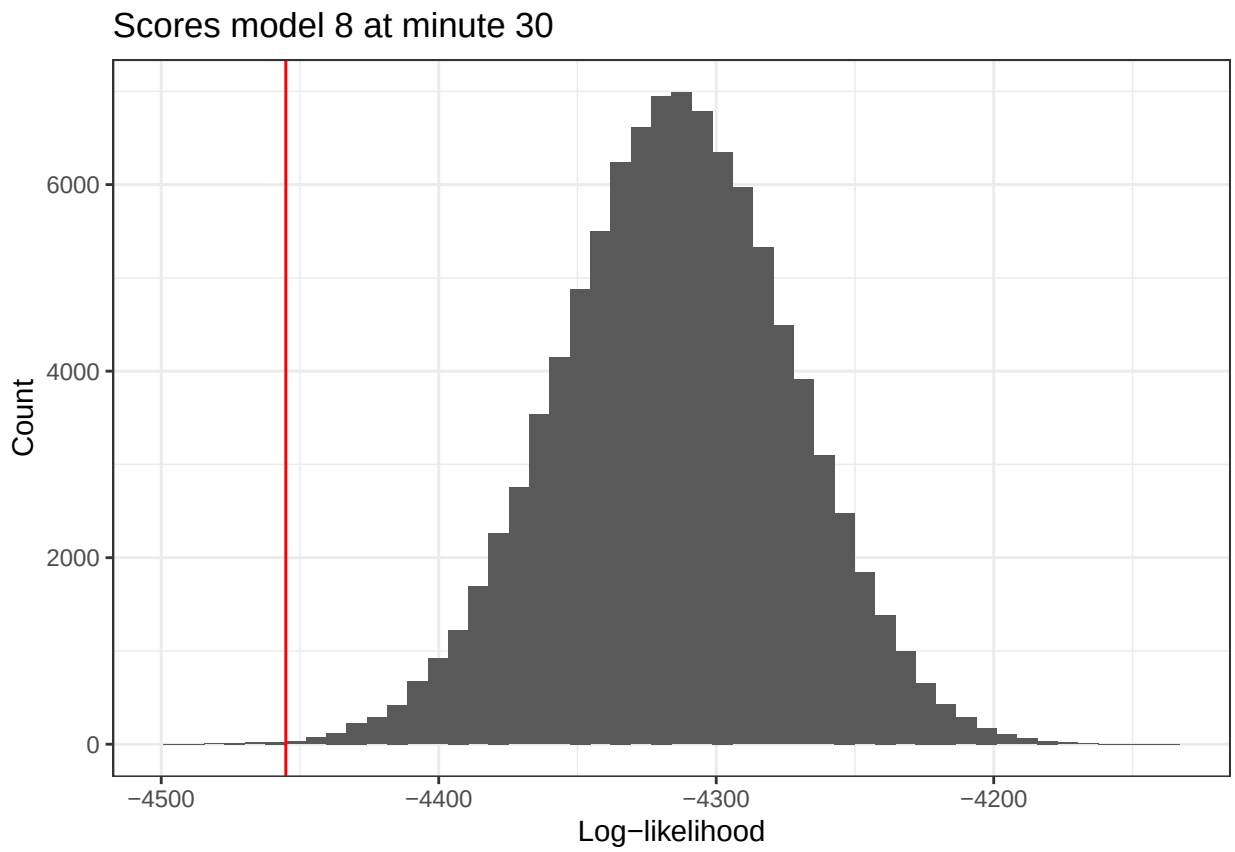
```
tibble(x = sims$pred_15$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_15, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 15")
```




```
sum(sims$pred_15$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_15) /
  length(sims$pred_15$loglik_scores_mod_8)
```

```
## [1] 0.00095
```

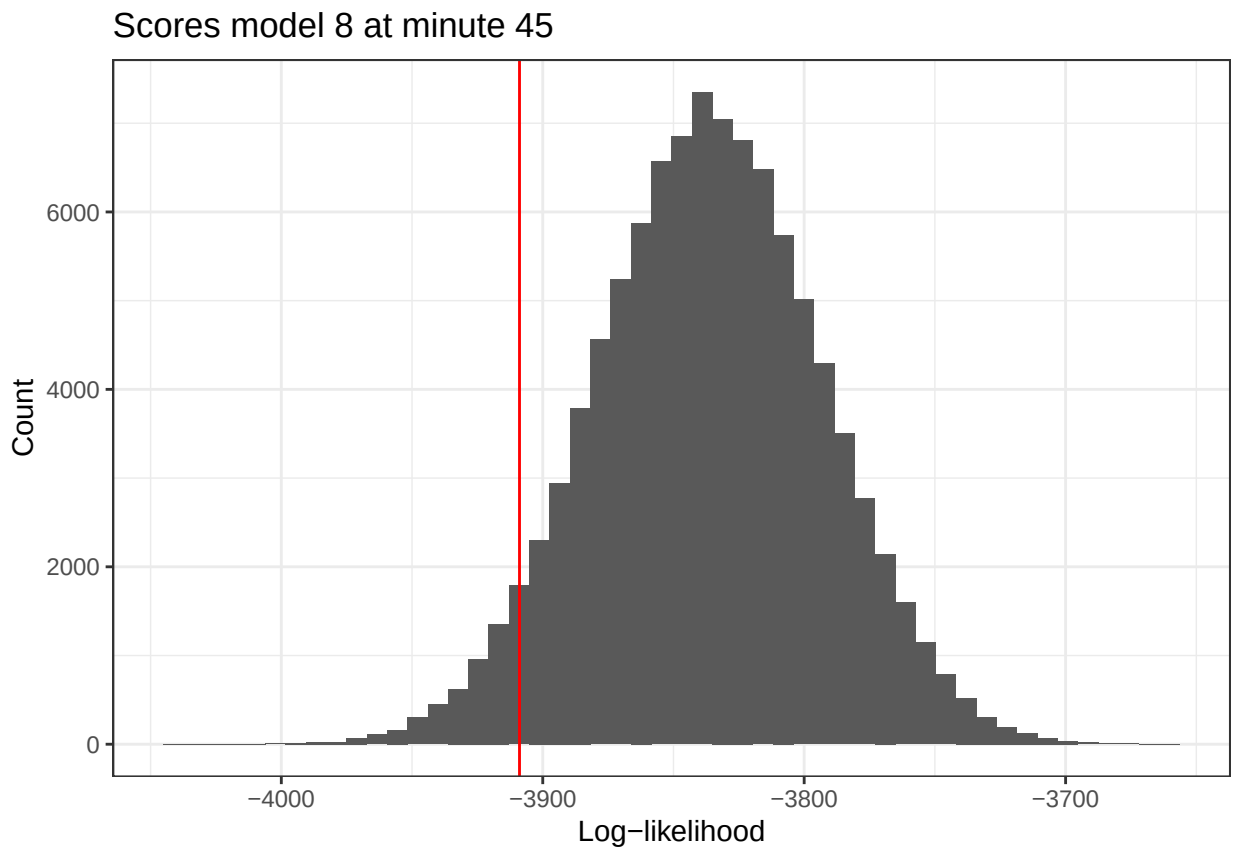
```
tibble(x = sims$pred_30$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_30, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 30")
```




```
sum(sims$pred_30$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_30) /
  length(sims$pred_30$loglik_scores_mod_8)
```

```
## [1] 6e-04
```

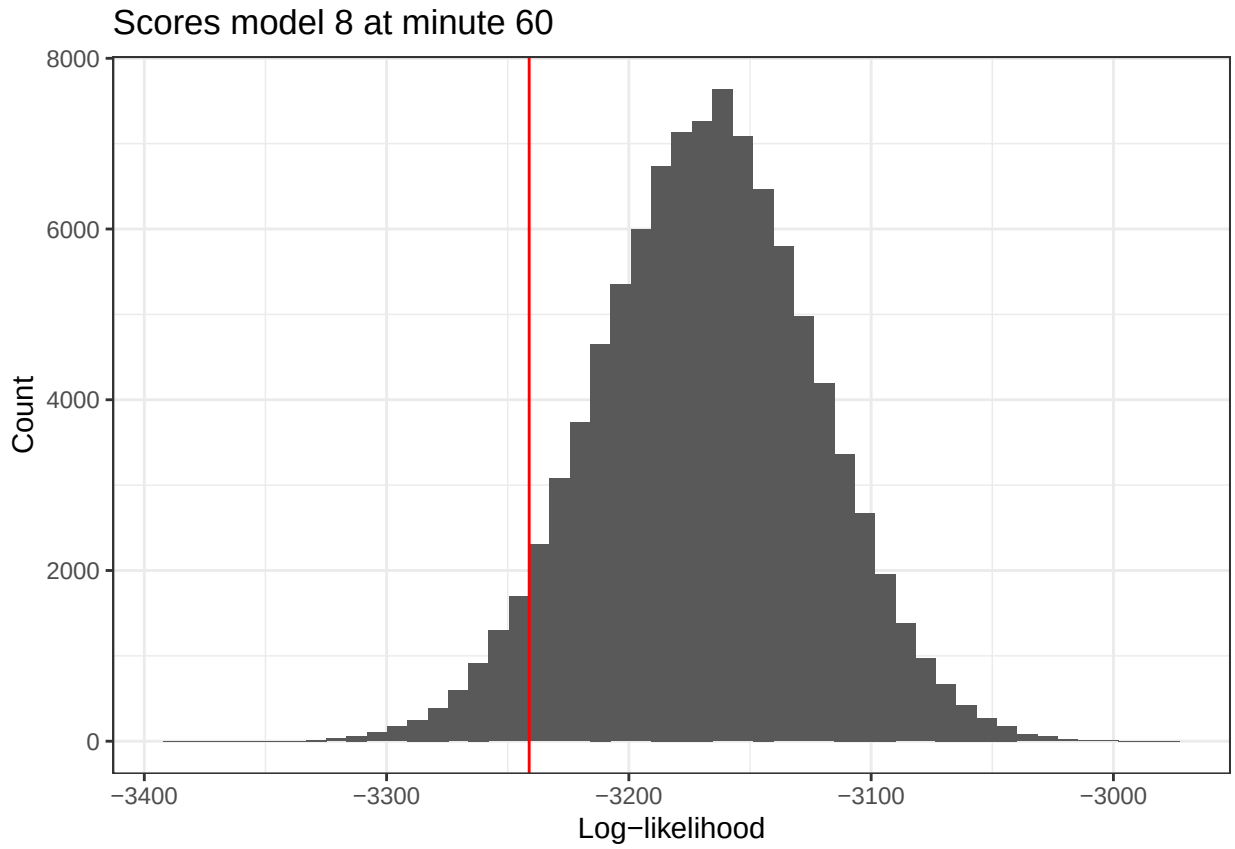
```
tibble(x = sims$pred_45$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_45, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 45")
```




```
sum(sims$pred_45$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_45) /
  length(sims$pred_45$loglik_scores_mod_8)
```

```
## [1] 0.04946
```

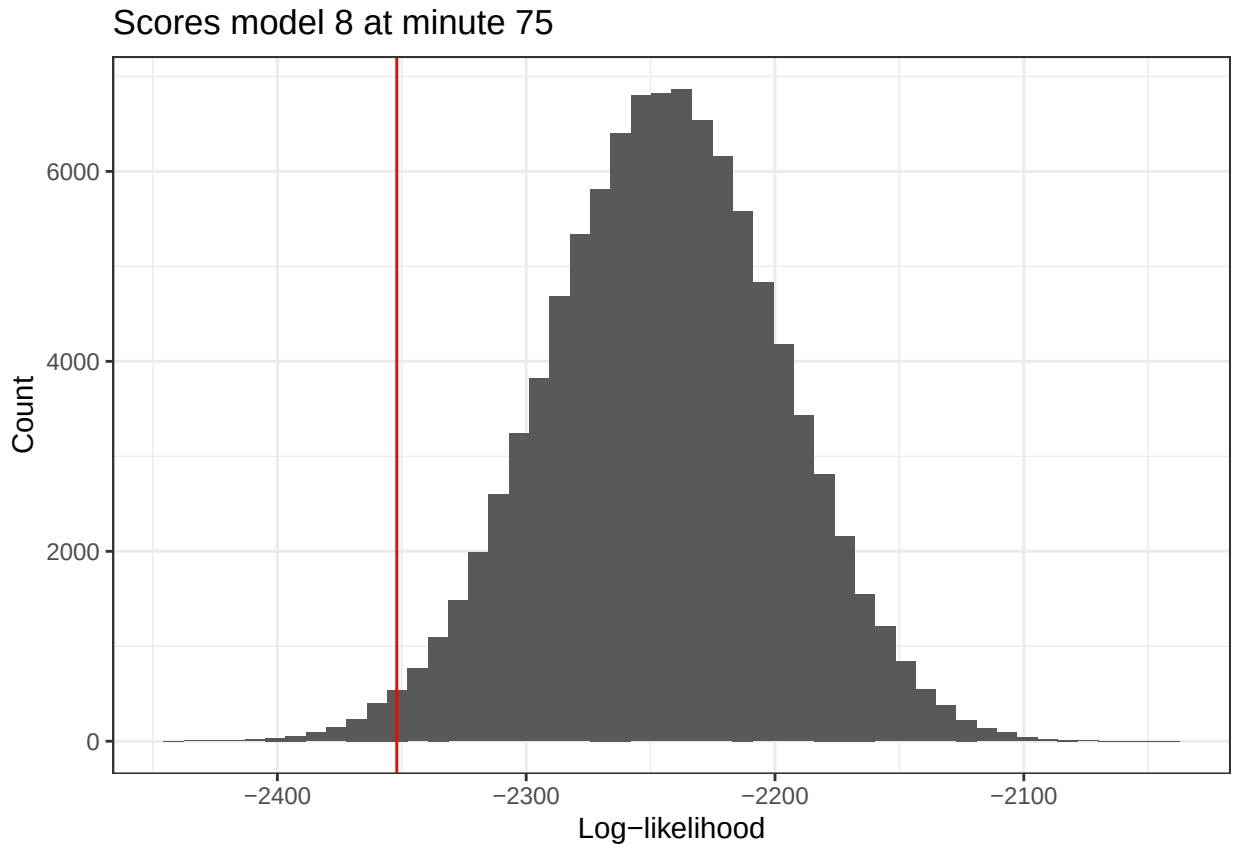
```
tibble(x = sims$pred_60$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_60, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 60")
```




```
sum(sims$pred_60$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_60) /
  length(sims$pred_60$loglik_scores_mod_8)
```

```
## [1] 0.05501
```

```
tibble(x = sims$pred_75$loglik_scores_mod_8) %>%
  ggplot(aes(x = x)) +
  geom_histogram(bins = 50) +
  theme_bw() +
  geom_vline(xintercept = loglik_observed_scores_mod_8_pred_75, col = "red") +
  xlab("Log-likelihood") +
  ylab("Count") +
  ggtitle("Scores model 8 at minute 75")
```




```
sum(sims$pred_75$loglik_scores_mod_8 <= loglik_observed_scores_mod_8_pred_75) /
  length(sims$pred_75$loglik_scores_mod_8)
```

```
## [1] 0.01235
```