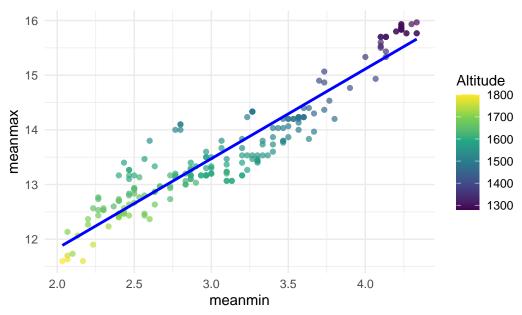
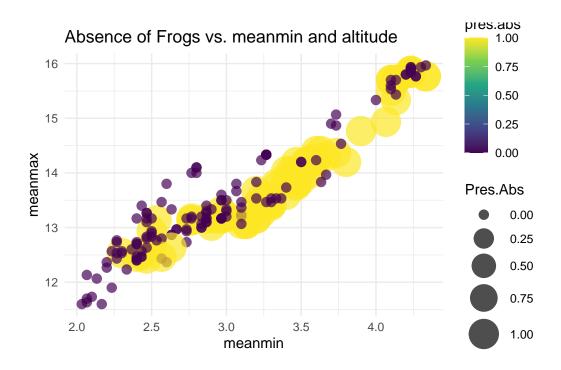
#### #Analysis of the data set

`geom\_smooth()` using formula = 'y ~ x'

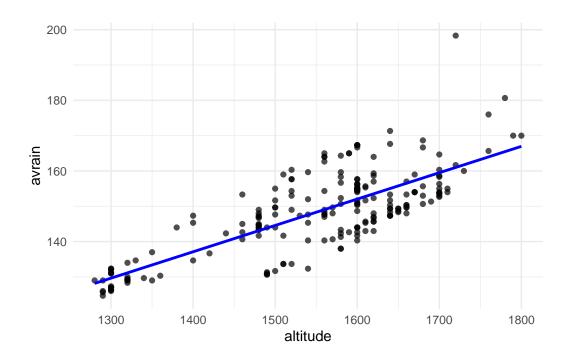
### Correlation between meanmin and meanmax with Altitude



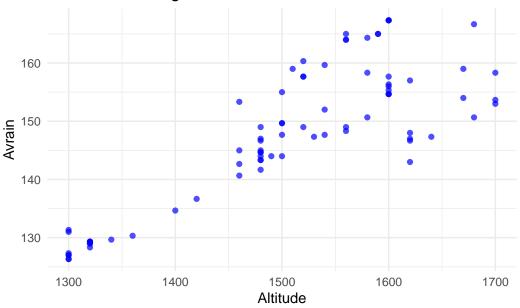


```
ggplot(data, aes(x = altitude, y = avrain)) +
  geom_point(alpha = 0.7) +
  geom_smooth(method = "lm", se = FALSE, color = "blue") + # Add linear regression line
  scale_color_viridis_c() + # You can change the color scale if needed
  theme_minimal()
```

`geom\_smooth()` using formula = 'y ~ x'



# Presence of Frogs vs. Altitude and Avrain

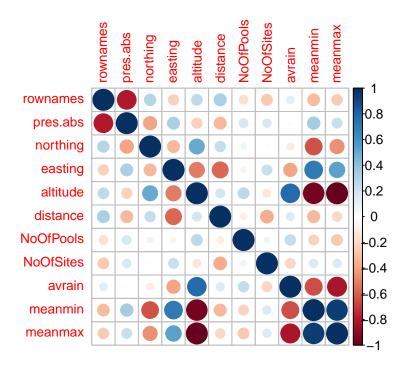


```
# Calculate correlations
correlation_matrix <- cor(data)
print(correlation_matrix)</pre>
```

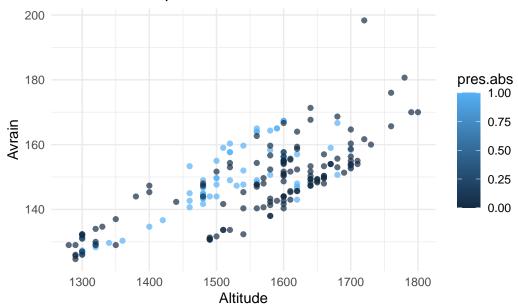
```
pres.abs
                                   northing
                                                easting
                                                          altitude
                                                                     distance
           rownames
           1.0000000 -0.79379062 0.29039960 -0.23733975
                                                         0.2594501
rownames
                                                                    0.3251945
pres.abs
         -0.7937906 1.00000000 -0.39875706 0.32493331 -0.2384100 -0.3217480
northing
          0.2903996 -0.39875706
                                 1.00000000 -0.32390138 0.5098837
                                                                    0.2416126
easting
          -0.2373398 0.32493331 -0.32390138 1.00000000 -0.5147038 -0.5753061
altitude
          0.2594501 -0.23841002 0.50988367 -0.51470383 1.0000000
                                                                    0.1820645
distance
          0.3251945 -0.32174803 0.24161260 -0.57530607
                                                         0.1820645
                                                                    1.0000000
NoOfPools -0.1727391 0.17562416 0.07394787 -0.06503599
                                                         0.2667913 -0.0959017
NoOfSites -0.2665277 0.16070989 0.01818349 0.22410207 -0.1368677 -0.3656936
          0.1257059 -0.01358179 -0.10914493 -0.39538028 0.7780271
avrain
          -0.3117578 0.34316473 -0.63320255 0.70890060 -0.9536610 -0.3263706
meanmin
meanmax
         -0.2505741 0.22265134 -0.45335004 0.53773711 -0.9965570 -0.2045681
           NoOfPools
                       NoOfSites
                                      avrain
                                                meanmin
                                                           meanmax
```

```
rownames -0.17273915 -0.26652773 0.12570585 -0.3117578 -0.2505741
         pres.abs
         0.07394787 \quad 0.01818349 \quad -0.10914493 \quad -0.6332025 \quad -0.4533500
northing
        -0.06503599 0.22410207 -0.39538028 0.7089006 0.5377371
easting
altitude
         0.26679134 -0.13686765 0.77802713 -0.9536610 -0.9965570
distance
        -0.09590170 -0.36569361 0.14848311 -0.3263706 -0.2045681
NoOfPools 1.00000000 0.07803560 0.24377919 -0.2237348 -0.2633172
NoOfSites 0.07803560 1.00000000 -0.21811450
                                        0.1510355 0.1576344
         0.24377919 -0.21811450 1.00000000 -0.6492240 -0.8186997
avrain
        -0.22373480 0.15103548 -0.64922402 1.0000000
                                                  0.9462741
meanmin
        1.0000000
meanmax
```

```
# Create a prettier plot of the correlation matrix
corrplot(correlation_matrix, method = "circle", type = "full", tl.cex = 0.8)
```



## Scatter Plot of pres.abs vs. Altitude and Avrain



## 1 Forming bayesian regression models using Stan

```
priors1 <- c(</pre>
  prior(normal(0,100), coef = "avrain"),
  prior(normal(0,100), coef = "NoOfPools"),
  prior(normal(15,10), coef = "meanmax"),
  prior(normal(4,3), coef = "meanmin"),
  prior(normal(1500,300), coef = "altitude")
fit1 <- brms::brm(</pre>
  # This specifies the formula
  pres.abs ~ altitude + avrain + NoOfPools + meanmin + meanmax,
  # This specifies the dataset
  data = data,
  # This specifies the observation model family
  family = "bernoulli",
  # This passes the priors specified above to brms
  prior = priors1,
  # This causes brms to cache the results
  file = "~/Documents/R/Project/fit1"
priors2 <- c(</pre>
  prior(normal(0,100), coef = "NoOfPools"),
  prior(normal(4,3), coef = "meanmin"),
```

```
prior(normal(1500,300), coef = "altitude")
fit2 <- brms::brm(</pre>
  # This specifies the formula
  pres.abs ~ NoOfPools + meanmin + altitude,
  # This specifies the dataset
  data = data,
  # This specifies the observation model family
  family = bernoulli(link="logit"),
  # This passes the priors specified above to brms
  prior = priors2,
  # This causes brms to cache the results
  file = "~/Documents/R/Project/fit2"
priors3 <- c(</pre>
 prior(normal(4,3), coef = "meanmin"),
  prior(normal(1500,300), coef = "altitude")
fit3 <- brms::brm(</pre>
  # This specifies the formula
  pres.abs ~ meanmin + altitude^2,
  # This specifies the dataset
  data = data,
  # This specifies the observation model family
  family = bernoulli(link="logit"),
 # This passes the priors specified above to brms
 prior = priors3,
  # This causes brms to cache the results
  file = "~/Documents/R/Project/fit3"
)
```

# 2 Posterior predictive checking and model analysis

```
pp_check_fit1 <- brms::pp_check(fit1, "stat")

Using all posterior draws for ppc type 'stat' by default.

traceplot_fit1 <- stanplot(fit1, type = "trace", prob= 0.95)

Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.</pre>
```

```
Warning: The following arguments were unrecognized and ignored: prob
No divergences to plot.
  pp_check_fit2 <- brms::pp_check(fit2,"stat")</pre>
Using all posterior draws for ppc type 'stat' by default.
  traceplot_fit2 <- stamplot(fit2, type = "trace")</pre>
Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.
No divergences to plot.
  pp_check_fit3 <- brms::pp_check(fit3,"stat")</pre>
Using all posterior draws for ppc type 'stat' by default.
  traceplot_fit3 <- stanplot(fit3, type = "trace")</pre>
Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.
No divergences to plot.
  s1 <- summary(fit1)</pre>
  s2 <- summary(fit2)</pre>
  s3 <- summary(fit3)
  print(s1)
 Family: bernoulli
  Links: mu = logit
Formula: pres.abs ~ altitude + avrain + NoOfPools + meanmin + meanmax
  Data: data (Number of observations: 212)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
Population-Level Effects:
          Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
                      118.16 -317.67 143.88 1.00
Intercept
            -78.25
                                                          1103
                                                                   1168
              0.03
                        0.03
                              -0.04
                                         0.10 1.00
                                                          1135
                                                                   1189
altitude
```

avrain	0.05	0.05	-0.06	0.16 1.00	1263	2004
NoOfPools	0.03	0.01	0.02	0.05 1.00	2634	2654
meanmin	6.14	1.36	3.52	8.92 1.00	2295	2448
meanmax	0.81	4.29	-7.26	9.41 1.00	1103	1328

Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS and Tail\_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

#### print(s2)

Family: bernoulli
Links: mu = logit

Formula: pres.abs ~ NoOfPools + meanmin + altitude

Data: data (Number of observations: 212)

Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup draws = 4000

#### Population-Level Effects:

	${\tt Estimate}$	${\tt Est.Error}$	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-60.90	11.79	-84.75	-39.50	1.00	1404	1870
NoOfPools	0.03	0.01	0.01	0.05	1.00	2383	2031
meanmin	6.72	1.16	4.55	9.08	1.00	1381	1881
altitude	0.02	0.01	0.01	0.04	1.00	1463	1894

Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS and Tail\_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

#### print(s3)

Family: bernoulli
Links: mu = logit

Formula: pres.abs ~ meanmin + altitude^2
Data: data (Number of observations: 212)

Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup draws = 4000

#### Population-Level Effects:

	Estimate	${\tt Est.Error}$	1-95% C	u-95%	CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-60.02	11.51	-82.7	-38	.71	1.00	1171	1309
meanmin	6.38	1.12	4.30	8	.61	1.00	1185	1419
altitude	0.03	0.01	0.0	2 0	.04	1.00	1191	1371

Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS and Tail\_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
r1 <- rhat(fit1)
  r2 <- rhat(fit2)
  r3 <- rhat(fit3)
  print(r1)
b_Intercept b_altitude
                           b_avrain b_NoOfPools
                                                   b_meanmin
                                                                b_{meanmax}
   1.001682
               1.001412
                            1.002236
                                        1.000579
                                                    1.001079
                                                                 1.001654
     lprior
                   lp__
   1.000519
               1.002178
  print(r2)
b_Intercept b_NoOfPools
                           b_meanmin b_altitude
                                                      lprior
                                                                     lp__
   1.004646
               1.000935
                            1.004281
                                        1.004620
                                                    1.004556
                                                                 1.001010
  print(r3)
b_Intercept
              b_meanmin b_altitude
                                          lprior
                                                         lp__
   1.002800
               1.002660
                            1.002540
                                        1.001230
                                                    1.000921
  mcse(fit1)
                      MCSE
    Parameter
1 b_Intercept 3.5691039928
2 b_altitude 0.0010223064
     b_avrain 0.0015271118
4 b_NoOfPools 0.0001690033
    b_meanmin 0.0285918130
    b_meanmax 0.1295099968
  mcse(fit2)
    Parameter
                      MCSE
1 b_Intercept 0.3127218690
2 b_NoOfPools 0.0001736244
    b_meanmin 0.0311892344
4 b_altitude 0.0001374450
  mcse(fit3)
    Parameter
                      MCSE
1 b_Intercept 0.3386960020
    b_meanmin 0.0327356654
3 b_altitude 0.0001516418
```

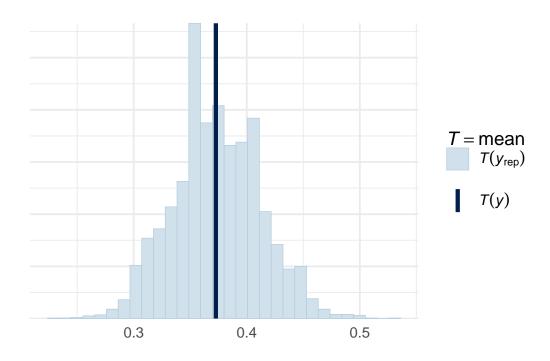
```
loo_fit1 <- loo(fit1)
loo_fit2 <- loo(fit2)
loo_fit3 <- loo(fit3)

loo_compare(loo_fit1,loo_fit2,loo_fit3)</pre>
```

```
elpd_diff se_diff
fit2 0.0 0.0
fit1 -1.3 1.4
fit3 -6.5 3.9
```

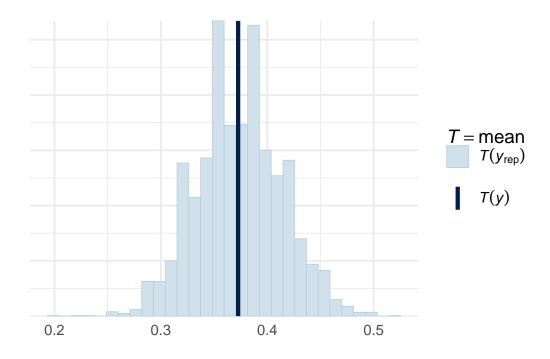
```
print(pp_check_fit1)
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



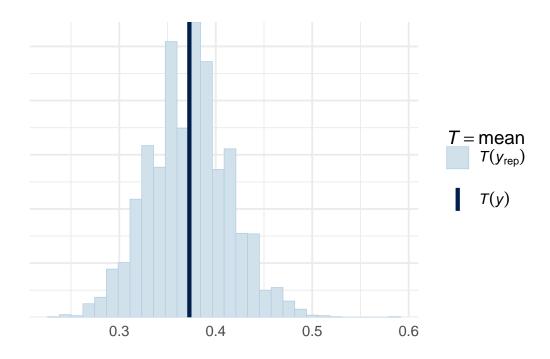
```
print(pp_check_fit2)
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



print(pp\_check\_fit3)

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



#Prior Sensitivity analysis

```
new_priors1_new <- c(</pre>
  prior(normal(0,100), coef = "avrain"),
  prior(normal(0,100), coef = "NoOfPools"),
  prior(normal(20,20), coef = "meanmax"),
  prior(normal(0,20), coef = "meanmin"),
  prior(normal(1000,1000), coef = "altitude")
fit_1_new_priors <- brms::brm(</pre>
  # This specifies the formula
  pres.abs ~ altitude + avrain + NoOfPools + meanmin + meanmax,
  # This specifies the dataset
  data = data,
  # This specifies the observation model family
  family = "bernoulli",
  # This passes the new_priors specified above to brms
  prior = new_priors1_new,
  # This causes brms to cache the results
  file = "~/Documents/R/Project/fit_1_new_priors"
)
new_priors2_new <- c(</pre>
  prior(normal(0,100), coef = "NoOfPools"),
  prior(normal(0,20), coef = "meanmin"),
 prior(normal(1000,1000), coef = "altitude")
)
fit_2_new_priors <- brms::brm(</pre>
  # This specifies the formula
  pres.abs ~ NoOfPools + meanmin + altitude,
  # This specifies the dataset
  data = data,
  # This specifies the observation model family
  family = bernoulli(link="logit"),
  # This passes the new_priors specified above to brms
  prior = new_priors2_new,
  # This causes brms to cache the results
  file = "~/Documents/R/Project/fit_2_new_priors_new_prior"
)
new_priors3_new <- c(</pre>
 prior(normal(0,20), coef = "meanmin"),
  prior(normal(1000,1000), coef = "altitude")
fit_3_new_priors <- brms::brm(</pre>
  # This specifies the formula
```

```
pres.abs ~ meanmin + altitude^2,
    # This specifies the dataset
    data = data,
    # This specifies the observation model family
    family = bernoulli(link="logit"),
    # This passes the new_priors specified above to brms
    prior = new_priors3_new,
    # This causes brms to cache the results
    file = "~/Documents/R/Project/fit_3_new_priors"
  pp_check_fit_1_new_priors <- brms::pp_check(fit_1_new_priors, "stat")</pre>
Using all posterior draws for ppc type 'stat' by default.
  traceplot_fit_1_new_priors <- stanplot(fit_1_new_priors, type = "trace", prob= 0.95)</pre>
Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.
Warning: The following arguments were unrecognized and ignored: prob
No divergences to plot.
  pp_check_fit_2_new_priors <- brms::pp_check(fit_2_new_priors, "stat")</pre>
Using all posterior draws for ppc type 'stat' by default.
  traceplot_fit_2_new_priors <- stanplot(fit_2_new_priors, type = "trace")</pre>
Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.
No divergences to plot.
  pp_check_fit_3_new_priors <- brms::pp_check(fit_3_new_priors, "stat")</pre>
Using all posterior draws for ppc type 'stat' by default.
  traceplot_fit_3_new_priors <- stanplot(fit_3_new_priors, type = "trace")</pre>
Warning: Method 'stanplot' is deprecated. Please use 'mcmc_plot' instead.
No divergences to plot.
```

```
s1_new <- summary(fit_1_new_priors)
s2_new <- summary(fit_2_new_priors)
s3_new <- summary(fit_3_new_priors)
print(s1_new)</pre>
```

Family: bernoulli
Links: mu = logit

Formula: pres.abs ~ altitude + avrain + NoOfPools + meanmin + meanmax

Data: data (Number of observations: 212)

Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup draws = 4000

#### Population-Level Effects:

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	-26.07	124.76	-264.68	220.92	1.01	1278	2048
altitude	0.01	0.04	-0.06	0.08	1.01	1314	2058
avrain	0.02	0.06	-0.09	0.13	1.01	1417	1994
NoOfPools	0.03	0.01	0.02	0.05	1.00	2312	2480
meanmin	6.66	1.49	3.87	9.67	1.00	2271	2420
meanmax	-1.28	4.58	-10.39	7.36	1.01	1282	1923

Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS and Tail\_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

### print(s2\_new)

Family: bernoulli
Links: mu = logit

Formula: pres.abs ~ NoOfPools + meanmin + altitude

Data: data (Number of observations: 212)

Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;

total post-warmup draws = 4000

#### Population-Level Effects:

	${\tt Estimate}$	Est.Error	1-95% CI	u-95% CI	Rhat	${\tt Bulk\_ESS}$	Tail_ESS
Intercept	-66.66	13.73	-93.98	-40.28	1.00	1630	2012
NoOfPools	0.03	0.01	0.01	0.05	1.00	2124	1691
meanmin	7.29	1.36	4.68	10.00	1.00	1631	1983
altitude	0.03	0.01	0.02	0.04	1.00	1660	2005

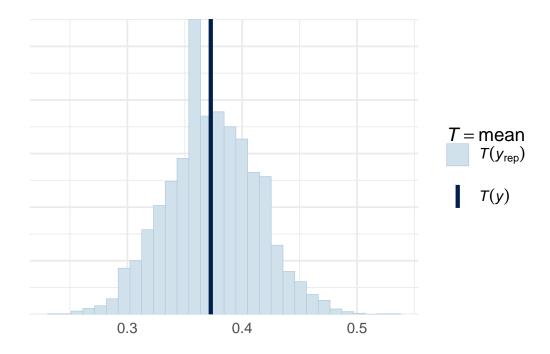
Draws were sampled using sample(hmc). For each parameter, Bulk\_ESS and Tail\_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

```
print(s3_new)
 Family: bernoulli
  Links: mu = logit
Formula: pres.abs ~ meanmin + altitude^2
   Data: data (Number of observations: 212)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
Population-Level Effects:
          Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
                        13.20
                                -90.41
                                         -39.12 1.01
Intercept
            -63.24
                                                           1215
                                                                    1407
                         1.29
                                  4.34
               6.70
                                            9.34 1.01
                                                           1202
                                                                    1416
meanmin
altitude
              0.03
                         0.01
                                  0.02
                                            0.04 1.01
                                                           1228
                                                                    1472
Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
  r1 <- rhat(fit_1_new_priors)</pre>
  r2_new <- rhat(fit_2_new_priors)</pre>
  r3_new <- rhat(fit_3_new_priors)</pre>
  print(r1)
b_Intercept b_altitude
                            b avrain b NoOfPools
                                                    b_meanmin
                                                                 b_meanmax
   1.006017
               1.005760
                            1.005071
                                         1.001669
                                                     1.003104
                                                                  1.006261
     lprior
                    lp__
   1.005584
               1.000081
  print(r2_new)
b_Intercept b_NoOfPools
                           b_meanmin b_altitude
                                                       lprior
                                                                      lp__
   1.001738
               1.001123
                            1.001881
                                         1.001626
                                                     1.002028
                                                                  1.000560
  print(r3_new)
b Intercept
              b_meanmin b_altitude
                                           lprior
                                                         lp__
   1.005434
               1.005309
                            1.005380
                                         1.004929
                                                     1.002796
  loo_fit_1_new_priors <- loo(fit_1_new_priors)</pre>
  loo_fit_2_new_priors <- loo(fit_2_new_priors)</pre>
  loo_fit_3_new_priors <- loo(fit_3_new_priors)</pre>
  loo_compare(loo_fit_1_new_priors,loo_fit_2_new_priors,loo_fit_3_new_priors)
```

```
elpd_diff se_diff
fit_2_new_priors 0.0 0.0
fit_1_new_priors -1.0 1.3
fit_3_new_priors -6.5 3.9
```

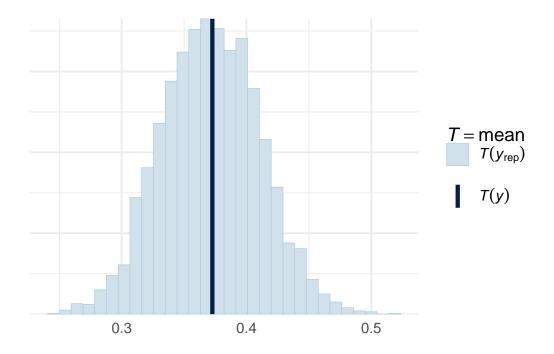
```
print(pp_check_fit_1_new_priors)
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



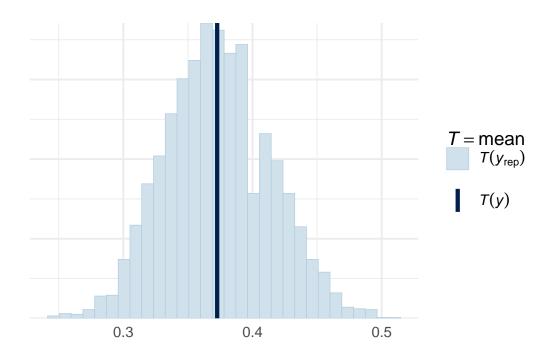
print(pp\_check\_fit\_2\_new\_priors)

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

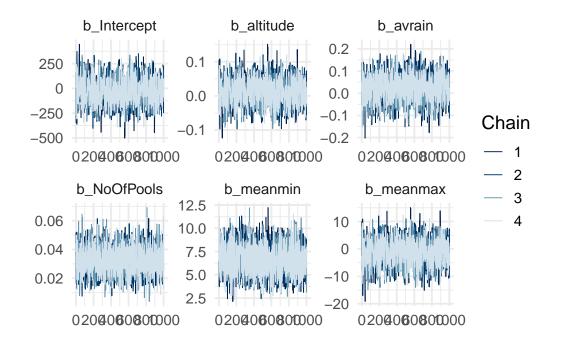


print(pp\_check\_fit\_3\_new\_priors)

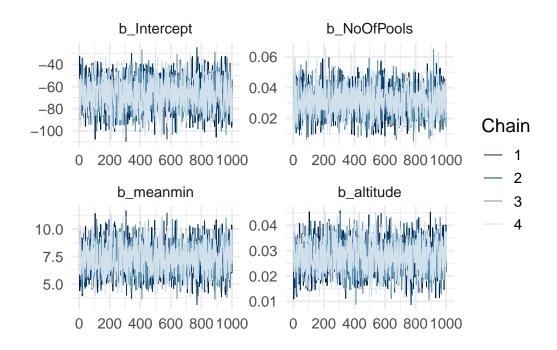
`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



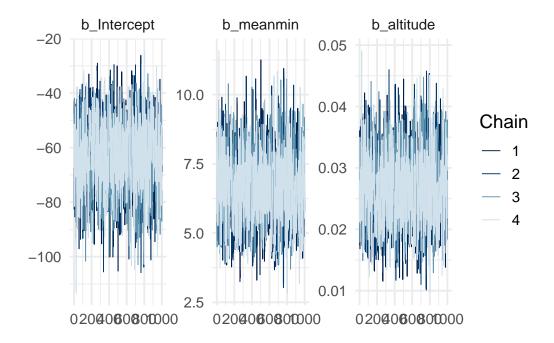
print(traceplot\_fit\_1\_new\_priors)



print(traceplot\_fit\_2\_new\_priors)



print(traceplot\_fit\_3\_new\_priors)



#### stancode(fit1)

```
// generated with brms 2.20.4
functions {
}
data {
  int<lower=1> N; // total number of observations
  array[N] int Y; // response variable
  int<lower=1> K; // number of population-level effects
  matrix[N, K] X; // population-level design matrix
  int<lower=1> Kc; // number of population-level effects after centering
  int prior_only; // should the likelihood be ignored?
}
transformed data {
  matrix[N, Kc] Xc; // centered version of X without an intercept
  vector[Kc] means_X; // column means of X before centering
  for (i in 2 : K) {
   means_X[i-1] = mean(X[:,i]);
    Xc[:, i-1] = X[:, i] - means_X[i-1];
  }
}
parameters {
  vector[Kc] b; // regression coefficients
  real Intercept; // temporary intercept for centered predictors
}
transformed parameters {
  real lprior = 0; // prior contributions to the log posterior
  lprior += normal_lpdf(b[1] | 1500, 300);
```

```
lprior += normal_lpdf(b[2] | 0, 100);
  lprior += normal_lpdf(b[3] | 0, 100);
  lprior += normal_lpdf(b[4] | 4, 3);
  lprior += normal_lpdf(b[5] | 15, 10);
  lprior += student_t_lpdf(Intercept | 3, 0, 2.5);
}
model {
  // likelihood including constants
  if (!prior_only) {
    target += bernoulli_logit_glm_lpmf(Y | Xc, Intercept, b);
  }
  // priors including constants
  target += lprior;
}
generated quantities {
  // actual population-level intercept
  real b_Intercept = Intercept - dot_product(means_X, b);
}
  stancode(fit2)
// generated with brms 2.20.4
functions {
}
data {
  int<lower=1> N; // total number of observations
  array[N] int Y; // response variable
  int<lower=1> K; // number of population-level effects
  matrix[N, K] X; // population-level design matrix
  int<lower=1> Kc; // number of population-level effects after centering
  int prior_only; // should the likelihood be ignored?
}
transformed data {
  matrix[N, Kc] Xc; // centered version of X without an intercept
  vector[Kc] means_X; // column means of X before centering
  for (i in 2 : K) {
    means_X[i-1] = mean(X[:,i]);
    Xc[:, i-1] = X[:, i] - means_X[i-1];
  }
}
parameters {
  vector[Kc] b; // regression coefficients
  real Intercept; // temporary intercept for centered predictors
}
transformed parameters {
  real lprior = 0; // prior contributions to the log posterior
  lprior += normal_lpdf(b[1] | 0, 100);
  lprior += normal_lpdf(b[2] | 4, 3);
```

```
lprior += normal_lpdf(b[3] | 1500, 300);
  lprior += student_t_lpdf(Intercept | 3, 0, 2.5);
}
model {
  // likelihood including constants
  if (!prior_only) {
    target += bernoulli_logit_glm_lpmf(Y | Xc, Intercept, b);
  // priors including constants
  target += lprior;
}
generated quantities {
  // actual population-level intercept
  real b_Intercept = Intercept - dot_product(means_X, b);
}
  stancode(fit3)
// generated with brms 2.20.4
functions {
}
data {
  int<lower=1> N; // total number of observations
  array[N] int Y; // response variable
  int<lower=1> K; // number of population-level effects
  matrix[N, K] X; // population-level design matrix
  int<lower=1> Kc; // number of population-level effects after centering
  int prior_only; // should the likelihood be ignored?
}
transformed data {
  matrix[N, Kc] Xc; // centered version of X without an intercept
  vector[Kc] means_X; // column means of X before centering
  for (i in 2 : K) {
    means_X[i - 1] = mean(X[:, i]);
    Xc[:, i-1] = X[:, i] - means_X[i-1];
  }
}
parameters {
  vector[Kc] b; // regression coefficients
  real Intercept; // temporary intercept for centered predictors
}
transformed parameters {
  real lprior = 0; // prior contributions to the log posterior
  lprior += normal lpdf(b[1] | 4, 3);
  lprior += normal_lpdf(b[2] | 1500, 300);
  lprior += student_t_lpdf(Intercept | 3, 0, 2.5);
}
model {
```

```
// likelihood including constants
if (!prior_only) {
   target += bernoulli_logit_glm_lpmf(Y | Xc, Intercept, b);
}
// priors including constants
  target += lprior;
}
generated quantities {
  // actual population-level intercept
  real b_Intercept = Intercept - dot_product(means_X, b);
}
```