1 Paddy Soil Adhesion (6 points possible)

We run the OpenBUGS code and obtain the results shown. The code is attached.

Node statistics

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
BR2	0.6748	0.07591	2.728E-4	0.4961	0.6857	0.7907	1001	100000
beta0	0.3991	0.1294	0.001104	0.1443	0.3981	0.6552	1001	100000
beta1	0.7682	0.08246	7.089E-4	0.605	0.7686	0.9304	1001	100000
pH.b0	0.9984	0.03972	1.777E-4	1.0	1.0	1.0	1001	100000
pH.b1	0.00292	0.05396	2.369E-4	0.0	0.0	0.0	1001	100000
rubber.pred	1.933	0.4227	0.00127	1.101	1.932	2.766	1001	100000

(a) (1 point) As the mean value of β_0 and β_1 is 0.3991 and 0.7682, respectively, the fitted model is

$$y = 0.3991 + 0.7682x,$$

where the response variable y is adhesion to rubber and the variable x is adhesion to steelB.

(1 point) The Bayesian R^2 is 0.6748. In python this value is 0.7133, so please do not penalize for small differences.

(b) (2 points) Any method students use to conclude that the model with $\beta_0 = 0$ and $\beta_1 = 1$ is not accurate is acceptable for credit. Solution 1: The credible set for β_0 is $[0.1443, 0.6552]$ and the credible set for β_1 is $[0.605, 0.9304]$. Since these do not contain 0 or 1 respectively, we conclude the model with $b_0 = 0$ and $b_1 = 1$ is not accurate.

Alternate solution: The first hypothesis we monitor is $P(\beta_0 > 0)$ using the `step()` function in BUGS. From the mean of the `pH.b0` row, we see that this probability is 0.9984, so we can be confident that β_0 is not equal to 0. Similarly for $P(\beta_1 > 1)$ we get a probability of 0.00292, thus we are confident that $\beta_1 < 1$. Other hypotheses and results are perfectly valid for credit as long as the student makes the correct conclusion.

(c) Based on the output, the predictive response has a mean of 1.933 and a 95% credible set as $[1.101, 2.766]$. Each of these values is worth 1 point.

Equivalent python output:

```
intercept  0.399  0.131    0.137  ...    4579.0    5402.0    1.0
beta       0.768  0.084    0.604  ...    4613.0    5034.0    1.0
sigma      0.418  0.049    0.328  ...    5978.0    5820.0    1.0
prob_H1    0.998  0.043    1.000  ...    5799.0   12000.0    1.0
prob_H2    0.004  0.060    0.000  ...    5405.0    5405.0    1.0

[5 rows x 9 columns]
r2        0.713324
r2_std    0.080090
dtype: float64
Prediction:

mean         1.936     | 100.00% [12000/12000 00:00<00:00]
sd           0.431
hdi_2.5%     1.115
hdi_97.5%    2.816
dtype: float64
```

2 Third-degree Burns (7 points possible)

By running the OpenBUGS code, we obtain the results shown as follows:

Node statistics

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta0	23.0	2.293	0.01346	18.72	22.94	27.67	1001	100000
beta1	-10.8	1.096	0.006418	-13.04	-10.77	-8.756	1001	100000
deviance	35.1	2.022	0.01307	33.15	34.48	40.48	1001	100000
p.star	0.8006	0.02652	1.308E-4	0.7467	0.8013	0.8505	1001	100000

(a) (2 points) Based on the estimated parameters, we obtain the logistic regression model as

$$y[i] \sim \mathcal{B}in(n_i, p_i), \text{ where } \mathrm{logit}(p_i) = 23.0 - 10.8x_i, i = 1, \dots, 9,$$

where y is the number of survivals, p is the probability of survival rate and n is total number patients.

(2 points) We obtain the following values for the deviance based on whether we model with a Bernoulli or Binomial likelihood:

Binomial + BUGS: 35.1

Binomial + pymc: 39.12

Bernoulli + BUGS: 337.2

Bernoulli + pymc: 339.43

(b) (3 points) When `log(area+1)` equals 2, the posterior probability of survival is 0.8006.
Equivalent python output for Bernoulli likelihood:

```
         mean     sd  hdi_3%  hdi_97%  ...  mcse_sd  ess_bulk  ess_tail  r_hat
beta0  22.934  2.301  18.497   27.286  ...    0.038    1835.0    1746.0    1.0
beta1 -10.768  1.099 -12.914   -8.722  ...    0.018    1834.0    1712.0    1.0

[2 rows x 9 columns]
Computed from 12000 posterior samples and 435 observations log-likelihood matrix.

              Estimate       SE
deviance_waic   339.43    24.50
p_waic            2.14        -

Predicted probability when log(area + 1) = 2:
 0.801
```

3 SO_2, NO_2, and Hospital Admissions (7 points possible)

By running the OpenBUGS code, we obtain the results shown as follows:

Node statistics

	mean	sd	MC_error	val2.5pc	median	val97.5pc	start	sample
beta0	5.458	0.01393	7.388E-4	5.411	5.46	5.473	1001	100000
beta1	-0.001859	3.248E-4	1.742E-5	-0.002358	-0.001831	-0.001432	1001	100000
beta2	0.002549	3.035E-4	1.657E-5	0.002251	0.002515	0.003267	1001	100000
lambda.star	279.1	3.003	0.1454	275.3	278.7	286.9	1001	100000
n.star	279.0	17.06	0.1587	246.0	279.0	313.0	1001	100000

(a) (2 points) We obtain the following model with the estimated parameters.

$$y_i \sim \mathcal{P}oi(\lambda_i), \text{ where } \lambda_i = \exp\{5.458 - 0.001859x_{i1} + 0.002549x_{i2}\}, i = 1, \ldots, n.$$

(2 points for correct interpretation) Based on the estimated coefficients, we see that the level of SO_2 has a negative coefficient (-0.001859) and the level of NO_2 has a positive coefficient (0.002549). This means that the level of SO_2 has a negaive impact on the hospital admission, while the level of NO_2 has a positive impact on the hospital admission.

(b) (3 points) Based on the results, the expected number of hospital admissions is 279.0 and the 95% credible set is $[246.0, 313.0]$.

Equivalent python output:

```
        mean     sd  hdi_2.5%  hdi_97.5%  ...  mcse_sd  ess_bulk  ess_tail  r_hat
beta0  5.461  0.007     5.448      5.476  ...      0.0    5152.0    5132.0    1.0
beta1 -0.002  0.000    -0.002     -0.001  ...      0.0    6359.0    5976.0    1.0
beta2  0.002  0.000     0.002      0.003  ...      0.0    4963.0    4933.0    1.0

[3 rows x 9 columns]
Sampling: [likelihood]
 |████████████████████| 100.00% [12000/12000 00:01<00:00]mean         278.457495
sd             16.787308
hdi_2.5%      244.686301
hdi_97.5%     309.894521
dtype: float64
```