

# Week 9: Hierarchical GLM

20/03/23

## Lip cancer

Here is the lip cancer data given to you in terribly unreproducible and error-prone format.

- `aff.i` is proportion of male population working outside in each region
- `observe.i` is observed deaths in each region
- `expect.i` is expected deaths, based on region-specific age distribution and national-level age-specific mortality rates.

### Question 1

Explain a bit more what the `expect.i` variable is. For example, if a particular area has an expected deaths of 6, what does this mean?

### Answer

Expected deaths is the implied number of lip cancer deaths for a particular region given that regions age structure and national level age-specific mortality rates for lip cancer. for example, and expected number of deaths would mean for that particular region we would expect 6 lip cancer deaths if this region were to experience the same age specific mortality rate as at the national level.

### Question 2

Run three different models in Stan with three different set-up's for estimating  $\theta_i$ , that is the relative risk of lip cancer in each region:

1. Intercept  $\alpha_i$  is same in each region  $= \alpha$  (with covariate)

2.  $\alpha_i$  is different in each region and modeled separately (with covariate)
3.  $\alpha_i$  is different in each region and the intercept is modeled hierarchically (with covariate)

### Answer

$$y_i | \theta_i \sim \text{Poisson}(\theta_i \cdot e_i)$$

look at three models for  $\log \theta_i$

Model 1:

$$\log \theta_i = \alpha + \beta x_i$$

Model 2:

$$\log \theta_i = \alpha_i + \beta x_i$$

Model 3:

$$\log \theta_i = \alpha_i + \beta x_i, \alpha_i \sim N(\mu, \sigma^2)$$

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 7e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.7 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 2000 [ 0%] (Warmup)

Chain 1: Iteration: 200 / 2000 [ 10%] (Warmup)

Chain 1: Iteration: 400 / 2000 [ 20%] (Warmup)

Chain 1: Iteration: 600 / 2000 [ 30%] (Warmup)

Chain 1: Iteration: 800 / 2000 [ 40%] (Warmup)

Chain 1: Iteration: 1000 / 2000 [ 50%] (Warmup)

Chain 1: Iteration: 1001 / 2000 [ 50%] (Sampling)

Chain 1: Iteration: 1200 / 2000 [ 60%] (Sampling)

Chain 1: Iteration: 1400 / 2000 [ 70%] (Sampling)

Chain 1: Iteration: 1600 / 2000 [ 80%] (Sampling)

Chain 1: Iteration: 1800 / 2000 [ 90%] (Sampling)

Chain 1: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 1:  
Chain 1: Elapsed Time: 0.102 seconds (Warm-up)  
Chain 1: 0.095 seconds (Sampling)  
Chain 1: 0.197 seconds (Total)  
Chain 1:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

Chain 2:  
Chain 2: Gradient evaluation took 2.1e-05 seconds  
Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.21 seconds.  
Chain 2: Adjust your expectations accordingly!  
Chain 2:  
Chain 2:  
Chain 2: Iteration: 1 / 2000 [ 0%] (Warmup)  
Chain 2: Iteration: 200 / 2000 [ 10%] (Warmup)  
Chain 2: Iteration: 400 / 2000 [ 20%] (Warmup)  
Chain 2: Iteration: 600 / 2000 [ 30%] (Warmup)  
Chain 2: Iteration: 800 / 2000 [ 40%] (Warmup)  
Chain 2: Iteration: 1000 / 2000 [ 50%] (Warmup)  
Chain 2: Iteration: 1001 / 2000 [ 50%] (Sampling)  
Chain 2: Iteration: 1200 / 2000 [ 60%] (Sampling)  
Chain 2: Iteration: 1400 / 2000 [ 70%] (Sampling)  
Chain 2: Iteration: 1600 / 2000 [ 80%] (Sampling)  
Chain 2: Iteration: 1800 / 2000 [ 90%] (Sampling)  
Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)  
Chain 2:  
Chain 2: Elapsed Time: 0.102 seconds (Warm-up)  
Chain 2: 0.076 seconds (Sampling)  
Chain 2: 0.178 seconds (Total)  
Chain 2:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

Chain 3:  
Chain 3: Gradient evaluation took 2.1e-05 seconds  
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.21 seconds.  
Chain 3: Adjust your expectations accordingly!  
Chain 3:  
Chain 3:  
Chain 3: Iteration: 1 / 2000 [ 0%] (Warmup)  
Chain 3: Iteration: 200 / 2000 [ 10%] (Warmup)  
Chain 3: Iteration: 400 / 2000 [ 20%] (Warmup)  
Chain 3: Iteration: 600 / 2000 [ 30%] (Warmup)  
Chain 3: Iteration: 800 / 2000 [ 40%] (Warmup)

```

Chain 3: Iteration: 1000 / 2000 [ 50%] (Warmup)
Chain 3: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 3: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 3: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 3: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 3: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 3:
Chain 3: Elapsed Time: 0.105 seconds (Warm-up)
Chain 3:                0.099 seconds (Sampling)
Chain 3:                0.204 seconds (Total)
Chain 3:

```

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

```

Chain 4:
Chain 4: Gradient evaluation took 2e-05 seconds
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.2 seconds.
Chain 4: Adjust your expectations accordingly!
Chain 4:
Chain 4:
Chain 4: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 4: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 4: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 4: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 4: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 4: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 4: Iteration:  1001 / 2000 [ 50%] (Sampling)
Chain 4: Iteration:  1200 / 2000 [ 60%] (Sampling)
Chain 4: Iteration:  1400 / 2000 [ 70%] (Sampling)
Chain 4: Iteration:  1600 / 2000 [ 80%] (Sampling)
Chain 4: Iteration:  1800 / 2000 [ 90%] (Sampling)
Chain 4: Iteration:  2000 / 2000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.122 seconds (Warm-up)
Chain 4:                0.093 seconds (Sampling)
Chain 4:                0.215 seconds (Total)
Chain 4:

```

	mean	se_mean	sd	2.5%	25%	50%
alpha	-0.008829978	0.0003504079	0.02118922	-0.04973236	-0.02297029	-0.008562268
beta	2.426033144	0.0029288241	0.17543942	2.08818765	2.30931138	2.425725585
	75%	97.5%	n_eff	Rhat		

alpha 0.004982719 0.03311239 3656.640 1.0014000  
beta 2.543136089 2.77441505 3588.127 0.9995667

## model 2

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

Chain 1:

Chain 1: Gradient evaluation took 5.6e-05 seconds

Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.56 seconds.

Chain 1: Adjust your expectations accordingly!

Chain 1:

Chain 1:

Chain 1: Iteration: 1 / 2000 [ 0%] (Warmup)

Chain 1: Iteration: 200 / 2000 [ 10%] (Warmup)

Chain 1: Iteration: 400 / 2000 [ 20%] (Warmup)

Chain 1: Iteration: 600 / 2000 [ 30%] (Warmup)

Chain 1: Iteration: 800 / 2000 [ 40%] (Warmup)

Chain 1: Iteration: 1000 / 2000 [ 50%] (Warmup)

Chain 1: Iteration: 1001 / 2000 [ 50%] (Sampling)

Chain 1: Iteration: 1200 / 2000 [ 60%] (Sampling)

Chain 1: Iteration: 1400 / 2000 [ 70%] (Sampling)

Chain 1: Iteration: 1600 / 2000 [ 80%] (Sampling)

Chain 1: Iteration: 1800 / 2000 [ 90%] (Sampling)

Chain 1: Iteration: 2000 / 2000 [100%] (Sampling)

Chain 1:

Chain 1: Elapsed Time: 0.479 seconds (Warm-up)

Chain 1: 0.458 seconds (Sampling)

Chain 1: 0.937 seconds (Total)

Chain 1:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

Chain 2:

Chain 2: Gradient evaluation took 2.3e-05 seconds

Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.23 seconds.

Chain 2: Adjust your expectations accordingly!

Chain 2:

Chain 2:

Chain 2: Iteration: 1 / 2000 [ 0%] (Warmup)

Chain 2: Iteration: 200 / 2000 [ 10%] (Warmup)

Chain 2: Iteration: 400 / 2000 [ 20%] (Warmup)

Chain 2: Iteration: 600 / 2000 [ 30%] (Warmup)

```

Chain 2: Iteration: 800 / 2000 [ 40%] (Warmup)
Chain 2: Iteration: 1000 / 2000 [ 50%] (Warmup)
Chain 2: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 2: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 2: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 2: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 2: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 2:
Chain 2: Elapsed Time: 0.513 seconds (Warm-up)
Chain 2: 0.584 seconds (Sampling)
Chain 2: 1.097 seconds (Total)
Chain 2:

```

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

```

Chain 3:
Chain 3: Gradient evaluation took 3.2e-05 seconds
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.32 seconds.
Chain 3: Adjust your expectations accordingly!
Chain 3:
Chain 3:
Chain 3: Iteration: 1 / 2000 [ 0%] (Warmup)
Chain 3: Iteration: 200 / 2000 [ 10%] (Warmup)
Chain 3: Iteration: 400 / 2000 [ 20%] (Warmup)
Chain 3: Iteration: 600 / 2000 [ 30%] (Warmup)
Chain 3: Iteration: 800 / 2000 [ 40%] (Warmup)
Chain 3: Iteration: 1000 / 2000 [ 50%] (Warmup)
Chain 3: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 3: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 3: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 3: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 3: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 3:
Chain 3: Elapsed Time: 0.602 seconds (Warm-up)
Chain 3: 0.582 seconds (Sampling)
Chain 3: 1.184 seconds (Total)
Chain 3:

```

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

```

Chain 4:
Chain 4: Gradient evaluation took 3.2e-05 seconds
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.32 seconds.

```

```

Chain 4: Adjust your expectations accordingly!
Chain 4:
Chain 4:
Chain 4: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 4: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 4: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 4: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 4: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 4: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 4: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 4: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 4: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 4: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 4: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.629 seconds (Warm-up)
Chain 4:                  0.458 seconds (Sampling)
Chain 4:                  1.087 seconds (Total)
Chain 4:

```

- Showing the first 5  $\alpha_i$  in the summary below:

	mean	se_mean	sd	2.5%	25%	50%
alpha[1]	-0.3325959	0.004914006	0.4112280	-1.18258514	-0.5984888	-0.3187823
alpha[2]	0.2824402	0.003191069	0.2792437	-0.31029092	0.1108496	0.2969458
alpha[3]	0.5083063	0.005640519	0.2675352	-0.01698017	0.3299046	0.5122248
alpha[4]	-0.3346987	0.005722626	0.4166076	-1.19927467	-0.6058283	-0.3112972
alpha[5]	0.5311704	0.005310633	0.3335934	-0.16672443	0.3095085	0.5473780
beta	1.4876979	0.021413908	0.5993505	0.27001258	1.0961057	1.4843950

  

	75%	97.5%	n_eff	Rhat
alpha[1]	-0.04272524	0.4342928	7003.1604	0.9996690
alpha[2]	0.47114152	0.8050398	7657.6322	0.9992894
alpha[3]	0.69456129	1.0090429	2249.6954	1.0011595
alpha[4]	-0.03199356	0.3962960	5299.8496	0.9997703
alpha[5]	0.76895237	1.1318605	3945.8656	0.9998689
beta	1.88610913	2.7118276	783.3754	1.0060433

### Model 3

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 1).

```

Chain 1:
Chain 1: Gradient evaluation took 8e-05 seconds
Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.8 seconds.
Chain 1: Adjust your expectations accordingly!
Chain 1:
Chain 1:
Chain 1: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 1: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 1: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 1: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 1: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 1: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 1: Iteration:  1001 / 2000 [ 50%] (Sampling)
Chain 1: Iteration:  1200 / 2000 [ 60%] (Sampling)
Chain 1: Iteration:  1400 / 2000 [ 70%] (Sampling)
Chain 1: Iteration:  1600 / 2000 [ 80%] (Sampling)
Chain 1: Iteration:  1800 / 2000 [ 90%] (Sampling)
Chain 1: Iteration:  2000 / 2000 [100%] (Sampling)
Chain 1:
Chain 1: Elapsed Time: 0.345 seconds (Warm-up)
Chain 1:                  0.301 seconds (Sampling)
Chain 1:                  0.646 seconds (Total)
Chain 1:

```

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 2).

```

Chain 2:
Chain 2: Gradient evaluation took 2.1e-05 seconds
Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.21 seconds.
Chain 2: Adjust your expectations accordingly!
Chain 2:
Chain 2:
Chain 2: Iteration:    1 / 2000 [  0%] (Warmup)
Chain 2: Iteration:   200 / 2000 [ 10%] (Warmup)
Chain 2: Iteration:   400 / 2000 [ 20%] (Warmup)
Chain 2: Iteration:   600 / 2000 [ 30%] (Warmup)
Chain 2: Iteration:   800 / 2000 [ 40%] (Warmup)
Chain 2: Iteration:  1000 / 2000 [ 50%] (Warmup)
Chain 2: Iteration:  1001 / 2000 [ 50%] (Sampling)
Chain 2: Iteration:  1200 / 2000 [ 60%] (Sampling)
Chain 2: Iteration:  1400 / 2000 [ 70%] (Sampling)
Chain 2: Iteration:  1600 / 2000 [ 80%] (Sampling)
Chain 2: Iteration:  1800 / 2000 [ 90%] (Sampling)
Chain 2: Iteration:  2000 / 2000 [100%] (Sampling)

```



Chain 2:  
Chain 2: Elapsed Time: 0.335 seconds (Warm-up)  
Chain 2: 0.298 seconds (Sampling)  
Chain 2: 0.633 seconds (Total)  
Chain 2:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 3).

Chain 3:  
Chain 3: Gradient evaluation took 2.1e-05 seconds  
Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.21 seconds.  
Chain 3: Adjust your expectations accordingly!  
Chain 3:  
Chain 3:  
Chain 3: Iteration: 1 / 2000 [ 0%] (Warmup)  
Chain 3: Iteration: 200 / 2000 [ 10%] (Warmup)  
Chain 3: Iteration: 400 / 2000 [ 20%] (Warmup)  
Chain 3: Iteration: 600 / 2000 [ 30%] (Warmup)  
Chain 3: Iteration: 800 / 2000 [ 40%] (Warmup)  
Chain 3: Iteration: 1000 / 2000 [ 50%] (Warmup)  
Chain 3: Iteration: 1001 / 2000 [ 50%] (Sampling)  
Chain 3: Iteration: 1200 / 2000 [ 60%] (Sampling)  
Chain 3: Iteration: 1400 / 2000 [ 70%] (Sampling)  
Chain 3: Iteration: 1600 / 2000 [ 80%] (Sampling)  
Chain 3: Iteration: 1800 / 2000 [ 90%] (Sampling)  
Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)  
Chain 3:  
Chain 3: Elapsed Time: 0.379 seconds (Warm-up)  
Chain 3: 0.375 seconds (Sampling)  
Chain 3: 0.754 seconds (Total)  
Chain 3:

SAMPLING FOR MODEL 'anon\_model' NOW (CHAIN 4).

Chain 4:  
Chain 4: Gradient evaluation took 2.6e-05 seconds  
Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.26 seconds.  
Chain 4: Adjust your expectations accordingly!  
Chain 4:  
Chain 4:  
Chain 4: Iteration: 1 / 2000 [ 0%] (Warmup)  
Chain 4: Iteration: 200 / 2000 [ 10%] (Warmup)  
Chain 4: Iteration: 400 / 2000 [ 20%] (Warmup)  
Chain 4: Iteration: 600 / 2000 [ 30%] (Warmup)  
Chain 4: Iteration: 800 / 2000 [ 40%] (Warmup)

```

Chain 4: Iteration: 1000 / 2000 [ 50%] (Warmup)
Chain 4: Iteration: 1001 / 2000 [ 50%] (Sampling)
Chain 4: Iteration: 1200 / 2000 [ 60%] (Sampling)
Chain 4: Iteration: 1400 / 2000 [ 70%] (Sampling)
Chain 4: Iteration: 1600 / 2000 [ 80%] (Sampling)
Chain 4: Iteration: 1800 / 2000 [ 90%] (Sampling)
Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)
Chain 4:
Chain 4: Elapsed Time: 0.408 seconds (Warm-up)
Chain 4:                0.375 seconds (Sampling)
Chain 4:                0.783 seconds (Total)
Chain 4:

```

	mean	se_mean	sd	2.5%	25%	50%
mu	0.08646907	0.0005342387	0.03673958	0.01404912	0.06194838	0.08683978
sigma	0.38753446	0.0006701095	0.03154214	0.32734433	0.36550779	0.38662842
beta	1.97464299	0.0065419806	0.33284476	1.31990404	1.75668818	1.97540626

  

	75%	97.5%	n_eff	Rhat
mu	0.1102819	0.1582667	4729.308	0.9996136
sigma	0.4085433	0.4511647	2215.597	1.0045886
beta	2.2007753	2.6363671	2588.600	1.0014371

### Question 3

Make two plots (appropriately labeled and described) that illustrate the differences in estimated  $\theta_i$ 's across regions and the differences in  $\theta$ s across models.

```

# A tibble: 6 x 4
      i median_mod1 lower_mod1 upper_mod1
  <int>    <dbl>    <dbl>    <dbl>
1     1      0.173      0.122      0.220
2     2      0.147      0.0987     0.193
3     3      0.556      0.466      0.646
4     4      0.309      0.247      0.370
5     5      0.379      0.309      0.447
6     6      0.398      0.326      0.469

```

```

# A tibble: 6 x 10
      i median~1 lower~2 upper~3 media~4 lower~5 upper~6 media~7 lower~8 upper~9
  <int>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
1     1      0.173      0.122      0.220    -0.208    -1.06      0.536      0.0221    -0.527      0.538

```

2	2	0.147	0.0987	0.193	0.391	-0.223	0.883	0.342	-0.110	0.772
3	3	0.556	0.466	0.646	0.863	0.394	1.29	0.799	0.366	1.20
4	4	0.309	0.247	0.370	-0.120	-0.987	0.581	0.113	-0.448	0.630
5	5	0.379	0.309	0.447	0.778	0.0793	1.35	0.652	0.0986	1.16
6	6	0.398	0.326	0.469	-0.463	-0.921	-0.0766	-0.275	-0.660	0.0685

# ... with abbreviated variable names 1: median\_mod1, 2: lower\_mod1,  
# 3: upper\_mod1, 4: median\_mod2, 5: lower\_mod2, 6: upper\_mod2,  
# 7: median\_mod3, 8: lower\_mod3, 9: upper\_mod3



