Problem 1

a) Rewriting the second expression for $f(y|\lambda)$ we obtain

$$f(y|\lambda) = e^{y\ln\lambda - \ln(e^\lambda - 1) - \ln y!} = e^{y\theta - b(\theta) - c(y)}$$

where $\theta = \ln\lambda$, $\lambda = e^\theta$ and $b(\theta) = \ln(e^{e^\theta} - 1)$.

b) Differentiating $b(\theta)$ we obtain

$$EY = \frac{e^{e^\theta}e^\theta}{e^{e^\theta} - 1} = \frac{\lambda e^\lambda}{e^\lambda - 1} = \frac{\lambda}{1 - e^{-\lambda}}$$

As $\lambda \to \infty$, the point mass $P(Y = 0) = e^{-\lambda}$ that we remove when truncating the distribution becomes negligable, so we should expect the mean to approach the mean if the distribution was not trucated, that is, λ.

This indeed happens as

$$\lim_{\lambda\to\infty} \frac{EY}{\lambda} = \lim_{\lambda\to\infty} \frac{1}{1 - e^{-\lambda}} = 1.$$

Problem 2

a) The model assumes that the number of covid cases

$$Y_i \sim \text{Poisson}(\mu_i)$$

where

$$\ln\mu_i = \beta_0 + \beta_1 \ln x_i$$

for at school $i = 1, 2, \ldots, n$ where x_i is the number of pupils at school i.

The parameter estimates are $\hat\beta_0 = -3.36$ and $\hat\beta_1 = 1.037$.

If we assume that the pupils at a particular school contract covid independently of each other with the same probability p_i, then Y_i would follow a binomial distribution. But because p_i is quite small it is reasonable to approximate this distribution with a Poisson distribution.

b) Under the null hypothesis of no overdispersion, the deviance D is chi-square with $n-p = 8-2 = 6$ degrees of freedom. We reject H_0 for large values of D. Hence the critical value is 12.59. Since the observed value is 14.28 we reject H_0 in favour of the alternative that ther is overdispersion.

An estimate of the overdisperion parameter φ is $\hat\varphi = D/(n-p) = 2.38$.

Adjusting for overdispersion, the standard errors of $\hat\beta_0$ and $\hat\beta_1$ change to $0.77\sqrt{2.38} = 1.1888$ and $0.1221\sqrt{2.38} = 0.1883$, respectively.

A possible mechanism would be that different pupils (within a particular school) do not contract covid-19 independently but with positive dependency. This would inflate the binomial variance (the variance of a sum of indicator variables) (which we have approximated by a Poisson variance).

c) The estimated relationship between the expected number of cases and the number of pupils is

$$E(Y_i) = e^{\beta_0+\beta_1 \ln x_i} = 0.03x_i^{1.034},$$

that is, almost direct proportionality between $E(Y_i)$ and x_i.

A possible simplification would be to build perfect proportionality into the model by instead including $\ln x_i$ as an offset by changing the linear predictor to

$$\ln \mu_i = \beta_0 + \ln x_i.$$

An interpretation of β_0 is that the expected rate of infection is

$$\frac{EY_i}{x_i} = e^{\beta_0}.$$

Relying on the standard error adjusted for overdispersion, we can test this null hypothesis of perfect proportionality against the fitted model using the Wald test stastic

$$T = \frac{\hat\beta_1 - 1}{\mathrm{SE}(\hat\beta_1)} = \frac{0.0368}{0.1883} = 0.1954.$$

which is smaller than the $t_{0.025,6} = 2.44$ in absolute value. Hence we do not reject H_0.

Problem 3

a) For child $j = 1, 2, \ldots, n_i$ of mother $i = 1, 2, \ldots, m$, we assume that the outcomes (birth delievered at hospital or elswhere), conditinal on the random mother effect γ_i, are independently distributed as

$$Y_{ij}|\gamma_i \sim \mathrm{Bernoulli}(\pi_i)$$

where

$$\operatorname{logit} \pi_{ij} = \beta_0 + \beta_1 \ln(\mathrm{income}_{ij}) + \beta_2 \mathrm{dist}_{ij} + \beta_3 \mathrm{dropout}_i + \beta_4 \mathrm{college}_i + \gamma_i$$

where $\mathrm{dropout}_i$ and $\mathrm{college}_i$ are 0/1-dummy variables indicating whether the ith mother dropped out of schools and finished collega education, respectively.

We also assume that the random mother effects $\gamma_1, \ldots, \gamma_m$ are iid $N(0, \tau^2)$.

The parameter estimates are $\hat\beta_0 = -3.29, \hat\beta_1 = 0.55, \ldots, \hat\beta_4 = 1.02$ and $\hat\tau^2 = 1.251$.

b) To test the signicance of the random intercept term, involves $H_0 : \tau^2 = 0$ versus $H_1 : \tau^2 > 0$. Since H_0 is on the boundary of the parameter space,

$$\text{LRT} = 2(l_1 - l_0)$$

where l_0 and l_1 are the maximum log-likelihoods under each alternative is distributed as a 50% : 50% mixture of chi-squares with 0 and 1 degree of freedom.

The critical value c satisfies

$$P(LRT > c) = \frac{1}{2}(\chi_0^2 > c) + \frac{1}{2}(\chi_0^2 > c) = \alpha$$

implying that $c = \chi^2_{1,2\alpha} = 6.63$.

Based on the observed value of

$$LRT = 2(-524.6 - (-537.5) = 25.8$$

we thus reject H_0 and conclude that there is evidence in the data for a difference between the different mothers.

c) For a given mother, that is, conditional on γ_i, we have

$$\text{logit}\, P(Y_{ij} = 1|\gamma_i) = \mathbf{x}_{ij}^T\boldsymbol{\beta} + \gamma_i.$$

Hence, the effect of of a unit change in college$_{ij}$, is that the log odds of birth delievery to a hospital changes by β_4, and the odds changes by a odds ratio of $e^{\beta_4} = e^{1.023} = 2.78$, that is, the odds increases by 178%. This interpretation is exact.

If we consider the average effect, we have the approximate relation

$$\begin{aligned} P(Y_{ij} = 1) &= \int P(Y_{ij} = 1|\gamma_i) f(\gamma_i) d\gamma_i \\ &= \int \frac{1}{1 + e^{-(\mathbf{x}_{ij}^T\boldsymbol{\beta} + \gamma_i)}} f(\gamma_i) d\gamma_i \\ &\approx \frac{1}{1 + e^{-\mathbf{x}_{ij}^T\boldsymbol{\beta}/\sqrt{1+.6\tau^2}}} \end{aligned}$$

or equivalently,

$$\text{logit}\, P(Y_{ij} = 1) \approx \mathbf{x}_{ij}^T\boldsymbol{\beta}/\sqrt{1 + .6\tau^2},$$

see Agresti (2002), p. 499. Thus, marginally, the odds of birth sdelivery at hospital changes approximately by a factor of $e^{1.023/\sqrt{1+0.6\cdot 1.25}} = 2.16$, that is, under the marginal model, the odds increases by 116% (as opposed to by 178% conditionally for a given mother).

Problem 4

a) Assuming equal strengths of both players, the probability of a white win is

$$P(Y_i = 1) = P(Y_i \le 1) = \Phi(\theta_1) = \Phi(-.2894) = 0.3861,$$

the probability of a draw is

$$P(Y_i = 2) = P(Y_i \le 2) - P(Y_i \le 1) = \Phi(\theta_2) - \Phi(\theta_1) = \Phi(0.8827) - \Phi(-.2894) = 0.4251,$$

and a black win has

$$P(Y_i = 3) = 1 - P(Y \le 2) = 1 - \Phi(\theta_2) = 0.1886.$$

b) The model corresponds to the assumption that the event $Y_i = r$ occurs if $\theta_{r-1} < u_i \le \theta_r$, $u_i = \alpha_{j(i)} - \alpha_{k(i)} + \varepsilon_i$, $\varepsilon_i \sim N(0,1)$ and $\theta_0 = -\infty$ and $\theta_3 = \infty$, since this implies $P(Y_i \le r) = \Phi(\theta_r + \alpha_{j(i)} - \alpha_{k(i)})$.

If a white and black win has equal probabilities (H_0),

$$\Phi(\theta_1) = 1 - \Phi(\theta_2).$$

The cdf of the standard normal has the property $\Phi(x) = 1 - \Phi(-x)$. Thus the null hypothesis corresponds to the linear hypothesis

$$\theta_1 = -\theta_2.$$

This can also be seen from the latent variable interpretation of the model, the latent variable

$$u_i = \underbrace{\alpha_{j(i)} - \alpha_{k(i)}}_{=0} + \underbrace{\varepsilon_i}_{\sim N(0,1)}$$

clearly falls below θ_1 (win to white) or above θ_2 (a win black) when θ_1 and θ_2 are located symmetrically around 0. This null hypothesis can be written on the form

$$\mathbf{C}\boldsymbol{\beta} = \mathbf{d}$$

where $\mathbf{C} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{d} = 0$.

To test this null hypothesis we use

$$\mathbf{C}\hat{\boldsymbol{\beta}} - \mathbf{d} = \hat{\theta}_1 + \hat{\theta}_2$$

as the starting point. This random variable has

$$\operatorname{Var}(\hat{\theta}_1 + \hat{\theta}_2) = \operatorname{Var}\hat{\theta}_1 + \operatorname{Var}\hat{\theta}_2 + 2\operatorname{Cov}(\hat{\theta}_1, \hat{\theta}_2)$$

Hence, relying on approximate/asymptotic normality of $\hat{\boldsymbol{\beta}}$,

$$Z = \frac{\hat{\theta}_1 + \hat{\theta}_2}{\sqrt{\text{Var}(\hat{\theta}_1 + \hat{\theta}_2)}}$$

is asymptotically/approximately $N(0,1)$ under H_0.

The observed value becomes

$$\frac{-.2894 + .8827}{\sqrt{0.0417 + 0.0528 + 2 \cdot 0.0176}} = 1.64$$

which is not in the critical region $|Z| > 1.96$ of a two sided test. Hence, the data do not provide sufficient evidence to conclude that there is any difference in the winning chances when playing with white versus black pieces.