#3 (a)

The density function for the Weibull is

$$f(x|\alpha, \lambda) = \alpha \lambda x^{\alpha-1} \exp(-\lambda x^\alpha)$$

with survival function

$$S(x|\alpha, \lambda) = \exp(-\lambda x^\alpha).$$

We retain 20,000 posterior samples after a burn-in of 5000. There didn't seem to be any issues regarding convergence.

I use four different sets of priors:

1. $p(\alpha) \propto 1/\alpha$, $p(\lambda) \propto 1/\lambda$

2. $\alpha \sim Ga(1,1)$, $\lambda \sim Ga(1,1)$

3. $\alpha \sim Ga(1, 1/10)$, $\lambda \sim Ga(1, 1/10)$

4. $\alpha \sim Ga(50, 50)$, $\lambda \sim Ga(50, 50)$

My preferred set is number 1, since this is non-informative while still not producing an improper posterior. Set 3 is also fairly non-informative, but is a proper prior. Sets 2 and 4 both give priors means to 1 for both parameters, but differ greatly in the prior variance (2 has much larger variance relative to 4).

From the figures (page 2), only prior set 4 gives much different results than the other 3, which is expected since set 4 is intended to be highly informative.

The posterior for α is each ploidy group is very similar across all prior sets (note the degree of overlap in the distributions). It may be reasonable to conclude that the value for α is common across the two groups.

The following table provides a numerical summary for the posteriors in each ploidy group under prior set 1.

Profile	Param	mean	var	hpd	
Aneuploid	α	0.827	0.015	0.584	1.066
	λ	0.114	0.001	0.047	0.191
Diploid	α	0.774	0.018	0.513	1.050
	λ	0.221	0.006	0.083	0.385

Table 1: Posterior summaries under prior set 1.
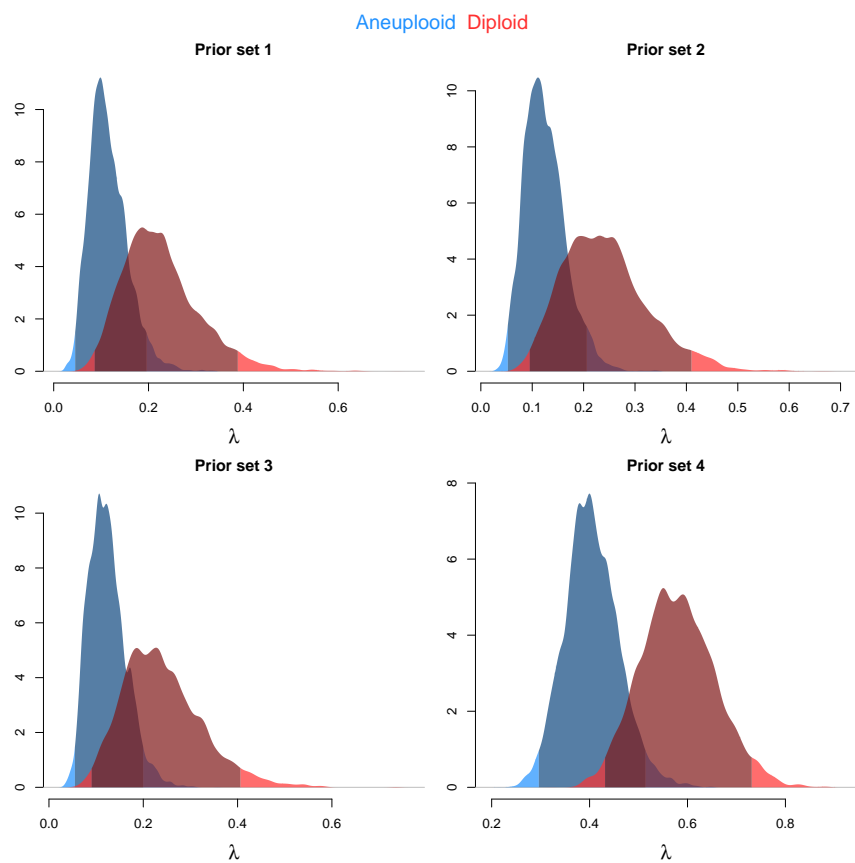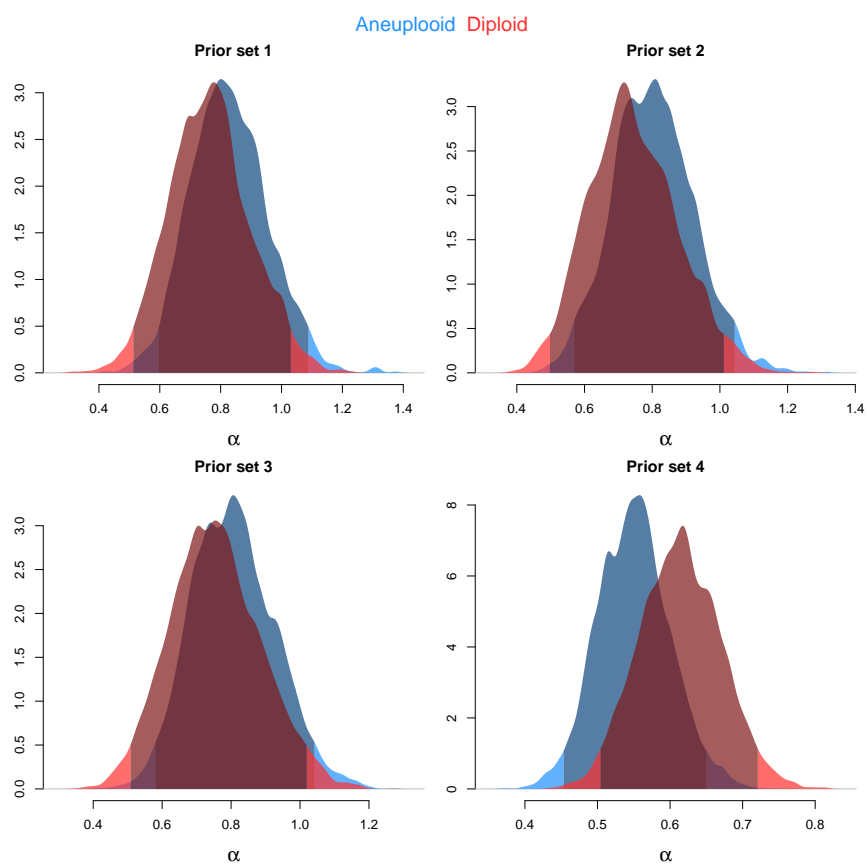
#3 (b-d)

I fit three models to the following:

$$Y_i = \log T_i = \beta_0 + \beta_1 x_i + \sigma G$$

where $x_i = 1$ if the individual is in the aneuploid group and $x_i = 0$ if in the diploid group, and G is one of three distributions: extreme value, normal, and logistic (all in standard form). These correspond to having T be distributed either Weibull, log-normal, and log-logistic, respectively.

The priors for each model are the same:

$$\beta_0 \sim N(0, 10^2), \qquad \beta_1 \sim N(0, 10^2), \qquad p(\sigma) \propto 1/\sigma.$$

These should allow the data to provide a greater contribution to the posteriors. Since I have no expert information to obtain informative priors, I tried to go as non-informative as possible.

After burning in 5000 samples, we keep 20000 posterior samples to be used in the analysis. There again did not seem to be any issue with convergence.

The following table provides means, variances, and hpds for the three parameters $(\beta_0, \beta_1, \sigma)$ in each of the three models. Plots of the distributions are shown on the next page. One thing to note as that for each model β_1 is very similar, whereas for the other parameters, this kind of similarity is not always seen.

Model	Param	mean	var	hpd		DIC
Weibull	β_0	2.022	0.078	1.484	2.589	
	β_1	0.689	0.141	-0.091	1.423	253.71
	σ	1.295	0.023	1.010	1.599	
Log-normal	β_0	1.364	0.121	0.675	2.065	
	β_1	0.800	0.188	-0.050	1.648	253.22
	σ	1.748	0.032	1.399	2.093	
Log-logistic	β_0	1.397	0.115	0.723	2.054	
	β_1	0.781	0.187	-0.097	1.615	253.45
	σ	1.000	0.014	0.776	1.247	

Table 2: Posterior summaries for the AFT model.

The coefficient β_1 tells us how to compare the two ploidy groups. Let $T(0)$ be a death time for someone in the diploid group (i.e. $x = 0$) and $T(1)$ be the time of death for someone in the aneuploid group $(x = 1)$. Then we have the following relation with the expected value of the death time

$$\begin{aligned}
E(T(0)) &= E(e^{\beta_0 + \beta_1 \times 0 + \sigma G}) \\
E(T(1)) &= E(e^{\beta_0 + \beta_1 \times 1 + \sigma G}) = e^{\beta_1} \times E(T(0))
\end{aligned}$$

For our Weibull model, we estimated $E(e^{\beta_1}) \approx 2.139$ (note: this is not $e^{0.689}$). This is to say we expect someone in the aneuploid group to live 2.139 times as long as someone in the diploid group.

DICs are calculated for each model (shown on the right side of the table). These are all very similar, so selecting which model to use is a matter of taste.