There are totally 38 points in the exam. The students with score higher than or equal to 35 points will receive 35 points. Please write down your name and student ID number below.

NAME: ________________________________
ID: _________________________________
1. (10 points) The table contains information of classroom behavior of 189 students. The students were classified on three factors: **rating**—teacher’s rating of classroom behavior (NonDeviant or Deviant), **risk**—risk index based on home conditions (NotRisk or Risk), and **adversity**—adversity of school conditions (Low, Medium, or High). R output is given.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Risk</th>
<th>Adversity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>NonDeviant</td>
<td>33(31.33)</td>
<td>29(30.27)</td>
</tr>
<tr>
<td></td>
<td>14(15.67)</td>
<td>67(65.73)</td>
</tr>
<tr>
<td>Deviant</td>
<td>3(4.67)</td>
<td>6(4.73)</td>
</tr>
<tr>
<td></td>
<td>4(2.33)</td>
<td>9(10.27)</td>
</tr>
</tbody>
</table>

> mod1
Call: glm(formula=yy~rating+risk+adversity,
family=poisson,data = behavior)

Degrees of Freedom: 11 Total (i.e. Null); 7 Residual
Null Deviance: 193.2
Residual Deviance: 31.45 AIC: 90.5

> mod2
Call: glm(formula=yy~(rating+risk+adversity)^2,
family = poisson,data = behavior)

Degrees of Freedom: 11 Total (i.e. Null); 2 Residual
Null Deviance: 193.2
Residual Deviance: 5.338 AIC: 74.39

> round(qchisq(0.95,1:20),2)

(a) (2 points). Test whether the three variables are independent.

(b) (2 points). Test whether the three-factor interaction effect is significant.
(c) (2 points). Compute the predict count at the low level of adversity with nonrisk for risk and nondeviant for rating under the main effect Poisson model.

(d) (2 points). Compute the predict count at the low level of adversity with nonrisk for risk and nondeviant for rating if the model contains all of the three main effect and risk-adversity interaction effect.

(e) (2 points). The predict counts of the model with rating-adversity and risk-adversity interaction effects are given in the table. Compute the residual deviance ($G^2$) and the residual Pearson $\chi^2$-statistic ($X^2$).
2. (10 points). The data set reported 10 different methods to fertilize tomatoes. Ten tomatoes of each method are weighted. The R output are given below.

```r
> summary(mod1)
Linear mixed-effects model fit by REML
Random effects:
  Formula: ~1 | method
     (Intercept) Residual
     StdDev: 15.68396 31.33399
Fixed effects: weight ~ 1
             Value Std.Error DF t-value p-value
(Intercept) 215.3375 5.866588 90 36.70575 0
```

```r
> summary(mod2)
Linear mixed-effects model fit by maximum likelihood
Random effects:
  Formula: ~1 | method
     (Intercept) Residual
     StdDev: 14.54544 31.33399
Fixed effects: weight ~ 1
             Value Std.Error DF t-value p-value
(Intercept) 215.3375 5.593572 90 38.49732 0
```

(a) (2 points). Describe the assumptions of the one-way random effect model in the output and provide estimates of model parameters in the REML and ML methods, respectively.

(b) (2 points). Describe the REML and ML methods.
(c) (2 points). Complete the following ANOVA table.

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>Fvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(d) (2 points). Provide the estimates of $\mu$ and $\sigma^2$ of the model: $y_{ij} \sim^i d N(\mu, \sigma^2)$.

(e) (2 points). Describe a way to test whether the random effect is significant.
3. (8 points). The data set reported routine information a few beef factories for supplying beef to market. The data set has five factories. Each factory provides information of 20 machines. Each machine is operated by 10 different (operators). Each operator cuts 10 slices of beef. The weights of these slices are measured. The target value of the weight is 500g.

```r
> summary(mod1)
Linear mixed-effects model fit by maximum likelihood
Random effects:
 Formula: ~1 | Factory
          (Intercept) StdDev: 17.89418
 Formula: ~1 | Machine %in% Factory
          (Intercept) StdDev: 1.708077
 Formula: ~1 | Operator %in% Machine %in% Factory
          (Intercept) Residual StdDev: 14.37122 5.031566
Fixed effects: Weight ~ 1
     Value Std.Error  DF t-value p-value
(Intercept) 496.979     8.017795 9000 61.9845 0
```

(a) (2 points). Provide the variance of the prediction of a new observed value.
(b) (2 points). Describe the nested model and provide estimates of model parameters.

(c) (2 points). Test whether the overall mean weight is significantly different from the target value.

(d) (2 points). Provide the 95% confidence interval for the new observed weight of a slice based on the output.
4. (10 point). The data reports the survival time in weeks for 40 patients with lung cancer with respect to two treatment (trt) methods.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Survival Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo (trt=1)</td>
<td>2 4 8+ 10 11 12+ 15 22+ 24+ 24 26 28 28+ 34 36+ 40 41 43+ 69 90</td>
</tr>
<tr>
<td>Treatment (trt=2)</td>
<td>3+ 6 8 10 12+ 14 20 26+ 26 79 84 90+ 120+ 136 141 149 165 166+ 187 198</td>
</tr>
</tbody>
</table>

> Diff
Call:
survdiff(formula=Surv(weeks,censor)~factor(trt))

<table>
<thead>
<tr>
<th>N</th>
<th>Observed</th>
<th>Expected</th>
<th>(0-E)^2/E</th>
<th>(O-E)^2/V</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor(trt)=1</td>
<td>20</td>
<td>13</td>
<td>8.15</td>
<td>2.89</td>
</tr>
<tr>
<td>factor(trt)=2</td>
<td>20</td>
<td>14</td>
<td>18.85</td>
<td>1.25</td>
</tr>
</tbody>
</table>

> summary(modw)
Call:
survreg(formula=Surv(weeks,censor)~factor(trt),dist="weibull")

<table>
<thead>
<tr>
<th>Value</th>
<th>Std. Error</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>3.754</td>
<td>0.234</td>
<td>16.06</td>
</tr>
<tr>
<td>factor(trt)2</td>
<td>1.021</td>
<td>0.325</td>
<td>3.14</td>
</tr>
<tr>
<td>Log(scale)</td>
<td>-0.172</td>
<td>0.158</td>
<td>-1.09</td>
</tr>
</tbody>
</table>

> summary(modcox)
Call:
coxph(formula = Surv(weeks, censor) ~ factor(trt))

| n= 40, number of events= 27 |
| coef | exp(coef) | se(coef) | z | Pr(>|z|) |
| factor(trt)2 | -1.0741 | 0.3416 | 0.4866 | -2.207 | 0.0273 * |

(a) (2 points). Compute the Kaplan-Meier estimator of the survival function of the placebo group for $t \leq 20$. Test whether the survival functions are all equal.
(b) (2 points). Estimate the hazard functions for both groups in the exponential model.

(c) (2 points). Estimate the hazard functions for both groups under the Weibull model. Test whether it can be reduced to an exponential model.

(d) (2 points). Complete the following table for estimates of survival functions from the Cox proportional hazard model.

<table>
<thead>
<tr>
<th>Month</th>
<th>Placebo</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.9634</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.9264</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.8893</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.8532</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.7850</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.7450</td>
<td></td>
</tr>
</tbody>
</table>

(e) (2 points). Provide a way to assess the assumption of proportional hazard property.