Homework 2

Tim Vigers

September 15, 2020

1 BD 1.1.1

1. Example (a)

   (a) Here let X be a R.V. indicating the diameter of a pebble and $Y = log(X)$. The logarithm of the diameter is normally distributed, so:
   $$P_Y(Y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{y-\mu}{\sigma})^2}$$
   To find the distribution of X, we can do a simple transformation using $\frac{d}{dx}Y = \frac{1}{X}$ and see that
   $$P_X(X) = \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{log(x)-\mu}{\sigma})^2}$$

   (b) Pebble diameters must be $X \in (0, \infty)$, so $-\infty < log(X) < \infty$. Because we are assuming $log(X) \sim \mathcal{N}(\mu, \sigma^2)$, $-\infty < \mu < \infty$ and $\sigma > 0$.

   (c) This is a parametric model because we are assuming a distribution for the pebble diameters.

2. Example (b)

   (a) For this example we have the model $X_i = \mu + \epsilon_i$, for $1 \leq i \leq n$ and $\epsilon \sim \mathcal{N}(0.1, \sigma^2)$. Therefore
   $$P_X(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu+0.1}{\sigma})^2}$$

   (b) In this case the variance of the errors is known, so the parameter space is $\mu \in R$.

(c) This is also a parametric model because we are assuming a distribution for the errors.

3. Example (c)

(a) This is similar to the model above, but this time $X_i = \mu + \epsilon_i$, for $1 \leq i \leq n$ and $\epsilon \sim \mathcal{N}(\theta, \sigma^2)$. Therefore

$$P_X(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu+\theta}{\sigma})^2}$$

(b) The variance of the errors is still known, but this time we are only able to estimate the parameter $\mu + \theta \in R$ as the model is unindentifiable for μ or θ alone.

(c) This is still a parametic model because we assume a distribution of the errors.

4. Example (d)

(a) Let $X =$ the number of eggs laid by an insect, which follows a Poisson distribution:

$$P_X(X) = \frac{e^{-\lambda}\lambda^x}{x!}$$

for $x = 0, 1, ...$ and $\lambda > 0$. If $Y =$ the number of eggs that hatch assuming each egg hatches with probability p, then Y follows a binomial distribution:

$$P_Y(Y|n = x) = \binom{x}{y} p^y (1-p)^{x-y}$$

(b)

$$\lambda > 0$$

$$Y = 0, 1, ...$$

$$0 \leq p \leq 1$$

(c) This is also a parametric model because we are assuming distributions for X and $Y|X$.

1.1 1.1.2

1. Problem 1.1.1(c): It is possible to estimate the parameter $\mu + \theta$, but it is not possible to estimate μ or θ separately because there are many possible values of μ and θ that would produce the same $\mu + \theta$ (for example ($\mu = 2, \theta = 2$) and ($\mu = 3, \theta = 1$)).

2. The parameterization of 1.1.1(d) is indentifiable because the entomologist is collecting the number of eggs laid by each insect, which allows for estimation of λ. They are also collecting the number of eggs hatching within a nest, which makes it possible to estimate p.

3. Unlike the case above, if the entomologist is only collecting data on the number of eggs hatched, they would only be able to estimate $Y|X$, but would not be able to estimate the average number of eggs laid per nest.