Pen and paper tasks assignment 1 - Philip Rettig

prettig@kth.se

3.1 Deriving the Log-Normal pdf using the transformation theorem

Gaussian PDF:
The Gaussian PDF for a random variable X is defined as:

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Transformation Function:
For the random variable Y:

$$Y = e^X$$

The inverse transformation is:

$$X = \ln(Y)$$

Derivative of x w.r.t y:

$$\frac{d}{dY} \ln(Y) = \frac{1}{Y}$$

Applying the Transformation Theorem:
The transformation theorem states that the PDF of Y, $p_Y(y)$, is given by:

$$p_Y(y) = p_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$\iff p_Y(y) = p_X(\ln(y)) \cdot \left| \frac{1}{y} \right|$$

Substituting the Gaussian PDF for $p_X(x)$ where $x = \ln(y)$:

$$p_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}} \cdot \frac{1}{y}$$

Answer:
The PDF of Y, which is the Log-Normal distribution, is:

$$p_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln(y)-\mu)^2}{2\sigma^2}}$$

Burglery

- Burglary every second year: $p(b=1) = \beta = \frac{1}{730}$ assuming we want the probability of a burglery happening on that day. If we want the probability of a burglery happening that year it is $p(b=1) = \beta = \frac{1}{2}$
- Earthquake every third year: $p(e=1) = \epsilon = \frac{1}{1095}$ assuming we want the probability of a earthquake happening on that day. If we want the probability of an earthquake happening that year it is $p(e=1) = \epsilon = \frac{1}{3}$ $p(b=1) = \beta = \frac{1}{3}$
- Alarm trigger due to burglary: $\alpha_b = 0.99$
- Alarm trigger due to earthquake: $\alpha_e = 0.01$
- Alarm trigger due to ghost/failure: $f = 0.001$

Probability of all possible values $p(a|b,e)$

- $p(a=1 \mid b=0, e=0) = f = 0.001$
- $p(a=0 \mid b=0, e=0) = 1 - f = 0.999$
- $p(a=1 \mid b=1, e=0) = \alpha_b = 0.99$
- $p(a=0 \mid b=1, e=0) = 1 - \alpha_b = 0.01$
- $p(a=1 \mid b=0, e=1) = \alpha_e = 0.01$
- $p(a=0 \mid b=0, e=1) = 1 - \alpha_e = 0.99$
- $p(a=0 \mid b=1, e=1) = (1-\alpha_b)(1-\alpha_e) = 0.0001$
- $p(a=1 \mid b=1, e=1) = 1 - (1-\alpha_b)(1-\alpha_e) = 0.9999$

Bayes' Theorem:

$$p(b=1|e=1, a=1) = \frac{p(a=1|b=1, e=1)p(b=1)p(e=1)}{p(a=1|e=1)}$$

Using the **law of total probability** for the denominator $p(a=1|e=1)$:

$$p(a=1|e=1) = p(a=1|b=1, e=1)p(b=1) + p(a=1|b=0, e=1)p(b=0)$$

$$p(b=0) = 1 - \frac{1}{730} = \frac{729}{730}$$

$$p(a=1|e=1) = 0.9901099 \cdot \frac{1}{730} + 0.01099 \cdot \frac{729}{730}$$

$$p(a=1|e=1) = 0.9901099 \cdot 0.00136986 + 0.01099 \cdot 0.99863014 = 0.001356 + 0.010964 = 0.01232$$

Applying Bayes' Theorem:

$$p(b=1|e=1,a=1) = \frac{p(a=1|b=1,e=1)p(b=1)p(e=1)}{p(a=1|e=1)}$$

$$p(b=1|e=1,a=1) = \frac{0.9999 \cdot \frac{1}{730} \cdot \frac{1}{1095}}{0.01232}$$

$$p(b=1|e=1,a=1) = \frac{0.999 \cdot 0.00000136986}{0.01232} \approx \frac{0.000001355}{0.01232} \approx 0.110$$

Posterior predictive

i)

$$p(y = k \mid D) = \int p(y = k \mid \theta)p(\theta \mid D)d\theta = \int \theta_k p(\theta_k \mid D)d\theta_k = E[\theta_k \mid D] = \frac{\hat{\alpha}_k}{\sum_{k'} \hat{\alpha}_{k'}}$$

So

$$p(y = k \mid \theta) = Cat(y \mid \theta), \theta \sim \text{Dir}(\alpha_1, \dots, \alpha_{27}), \quad \text{where } \alpha_k = 10 \text{ for all } k$$

$$p(x_{2001} = e \mid D) = \int p(x_{2001} = e \mid \theta)p(\theta \mid D)d\theta = E[\theta_k \mid D] = \frac{\hat{\alpha}_k}{\sum_{k'} \hat{\alpha}_{k'}}$$

Which is the predictive distribution of a new observation given the updated Dirichlet.

Denominator:

$$\sum_{k'} \hat{\alpha}_{k'} = 270 + 2000 = 2270$$

Numerator:

$$\alpha'_e = \alpha_e + 260 = 10 + 260 = 270$$

$$p(x_{2001} = e \mid D) = \frac{270}{2270} = 0.11894$$

ii) Updated values

$$\alpha_e = \alpha_e + 260 = 10 + 260 = 270$$

$$\alpha_a = \alpha_a + 100 = 10 + 100 = 110$$

$$\alpha_p = \alpha_p + 100 = 10 + 87 = 97$$

$$\sum_{k'} \hat{\alpha}_{k'} = 270 + 100 + 97 + 2000 = 2467$$

$$p(x_{2001} = e \mid D) = \frac{270}{2467}$$

$$p(x_{2002} = a \mid D) = \frac{100}{2467}$$

$$p(x_{2001} = e, x_{2002} = a \mid D) = p(x_{2001} = e \mid D) \cdot (x_{2002} = a \mid D) = \frac{270}{2467} \cdot \frac{100}{2467} = 0.0044$$