Stat 8003, Homework 3

Group G: Andrew Schneider, Abdulsalam Hdadi, David Dobor

September 17, 2014

Question 3.1. Consider a bivariate distribution with $P(X = 1, Y = 2) = 0.4, P(X = 2, Y = 3) = 0.6$. Find the correlation coefficient between X and Y.

Answer: We fill the following table with the given joint probability density. (If X and Y attain other values, the joint density at those values must be zero, since the given probabilities already sum up to 1.) The marginal probabilities are given at the right and at the bottom margins of the table. All entries in the joint table sum up to 1; the entries in the margins also sum up to 1:

	$X = 1$	$X = 2$	$P(Y = y)$:
$Y = 2$	2/5	0	*2/5*
$Y = 3$	0	3/5	*3/5*
$P(X = x)$:	*2/5*	*3/5*	***1***

The correlation coefficient here is 1. We verify this by direct computation:

$$E[X] = 1 \cdot \frac{2}{5} + 2 \cdot \frac{3}{5} = \frac{8}{5}$$

$$E[X^2] = 1^2 \cdot \frac{2}{5} + 2^2 \cdot \frac{3}{5} = \frac{14}{5}$$

$$Var(X) = E[X^2] - (E[X])^2 = \frac{14}{5} - \left(\frac{8}{5}\right)^2 = \frac{6}{25}$$

$$E[Y] = 2 \cdot \frac{2}{5} + 3 \cdot \frac{3}{5} = \frac{13}{5}$$

$$E[Y^2] = 2^2 \cdot \frac{2}{5} + 3^2 \cdot \frac{3}{5} = 7$$

$$Var(Y) = E[Y^2] - (E[X])^2 = 7 - \left(\frac{13}{5}\right)^2 = \frac{6}{25}$$

$$E[XY] = 1 \cdot 2 \cdot \frac{2}{5} + 2 \cdot 3 \cdot \frac{3}{5} = \frac{22}{5}$$
$$E[X] \cdot E[Y] = \frac{8}{5} \cdot \frac{13}{5} = \frac{104}{25}$$
$$Cov(X,Y) = E[XY] - E[X] \cdot E[Y] = \frac{22}{5} - \frac{104}{25} = \frac{6}{25}$$

Therefore, the correlation coefficient is:

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}} = \frac{6/25}{\sqrt{6/25 \cdot 6/25}} = 1$$

Question 3.2. Find two random variables X and Y, such that $Cov(X,Y) = 0$ but X and Y are not independent.

Answer: Let X be Bernoulli:

$$X = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ +1 & \text{with probability } \frac{1}{2} \end{cases}$$

And Let Y be a `r.v.` that directly depends on the outcome of X as follows:

$$Y = \begin{cases} 0 & \text{whenever } \; X = -1 \\ -8003 & \text{with probability } \frac{1}{2} \; \text{ whenever } \; X = 1 \\ +8003 & \text{with probability } \frac{1}{2} \; \text{ whenever } \; X = 1 \end{cases}$$

Now since

$$\begin{aligned} E[XY] &= (-1) \cdot 0 \cdot P(X = -1) + \\ &\quad + 1 \cdot 8003 \cdot P(X = 1, Y = 1) + \\ &\quad + 1 \cdot (-8003) \cdot P(X = 1, Y = -1) \\ &= 0, \end{aligned}$$

and since $E[X]$ is also 0, the the covariance $Cov(X,Y) = E[XY] - E[X] \cdot E[Y] = 0$, although the `r.v.`s are by definition dependent.

Question 3.3. In the Example of `GDP`. Assume that the data follows a gamma distribution $\Gamma(\alpha, \beta)$.

(a) Derive the estimator for α, β using the methods of moments;

(b) Compare the density of the data vs the fitted curve.

Answer:

(a) We use the following paametrization of Gamma, where α is called the *shape* parameter and β is called the *scale* parameter:

$$f(x) = \frac{1}{\Gamma(\alpha)} \cdot \frac{1}{\beta} \cdot \left(\frac{x}{\beta}\right)^{\alpha-1} \mathrm{e}^{-\frac{x}{\beta}} \qquad \text{where } \ x \geq 0 \ \ \alpha, \beta > 0$$

Since $\mathrm{E}\, X = \alpha\beta$ and $\mathrm{E}\, X^2 = \alpha(\alpha+1)\beta^2$, we compute the first and second sample moments from the data (denoting them by m_1 and m_2, respectively) and set them equal to the distribution parameters as follows:

$$\begin{cases} m_1 = \hat{\alpha}\hat{\beta} \\ m_2 = \hat{\alpha}(\hat{\alpha}+1)\hat{\beta}^2 = \hat{\alpha}^2\hat{\beta}^2 + \hat{\alpha}\hat{\beta}^2 \end{cases} \Rightarrow \begin{cases} m_1 = \hat{\alpha}\hat{\beta} \\ m_2 - m_1^2 = \hat{\alpha}\hat{\beta}^2 \end{cases} \Rightarrow \begin{cases} \hat{\beta} = \frac{m_2 - m_1^2}{m_1} \\ \hat{\alpha} = \frac{m_1^2}{m_2 - m_1^2} \end{cases}$$

(b) We will use the estimates derved in part (a) to plot both the density and the fitted curve. We will pass to `R` the value obtained for $\hat{\alpha}$ as `shape` parameter, and the value obtained for $\hat{\beta}$ as the `scale` parameter:

```
m1 <- mean(gdp)
m2 <- mean(gdp^2)

beta.hat <- (m2 - m1^2) / m1
alpha.hat <- m1 / beta.hat

points(x, dgamma(x, shape=alpha.hat, scale=beta.hat), 'l', col='red')
```

....TO INSERT FIGURE HERE ....

Question 3.4. For any random variable X, let $M_X(t) = \mathrm{E}\, \mathrm{e}^{X(t)}$ and $S_X(t) = log(M_X(t))$. $M_X(t)$ is called the moment generating function and $SX(t)$ is the cumulant generating function. It is known that (Can you prove it? Not required.)

$$\frac{d}{dt} S_X(t)\Big|_{t=0} = \mathrm{E}\, X \quad \text{and} \quad \frac{d^2}{dt^2} S_X(t)\Big|_{t=0} = Var(X)$$

Use this fact to answer the following questions.

(a) Assume that X follows a Gamma distribution with parameter α and β. Calculate the cumulant generating function of $\log(X)$;

(b) Calculate $\mathrm{E}\ \log(X)$ and $Var(\log X)$. Write your final result by using the digamma function $\psi(x)$ and trigamma function $\psi_1(x)$, where $\psi(x) = (\log\Gamma(x))'$ and $\psi_1(x) = (\log\Gamma(x))''$.

(c) Match the first and second moment of $\log(X)$, and derive the `MOM` estimator of α and β.
(Hint: in `R`, you can use `digamma(x)`, `trigamma(x)`, and `limma::trigammaInverse(x)`.)

(d) Apply your estimator to the `GDP` dataset and estimate the parameters α and β.

Answer:

Let $Y = \log(X)$ where $X \sim \Gamma(\alpha, \beta)$. The distribution of $Y \sim$ `Log-of-Gamma`(α, β) is given by

$$f_Y(y) = \frac{1}{\beta^\alpha \Gamma(\alpha)}\ \mathrm{e}^{\left(\alpha y - e^{y/\beta}\right)}\, 1_{(-\infty,\infty)}(y)$$

Before answering part (a), we first find the moment generating function of `Log-of-Gamma`(α, β).

Moment Generating Function of Log Of Gamma

We now find the Moment Generating Function (`MGF`) of `Log-Of-Gamma`(α, β). By definition:

$$M_Y(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{ty} \frac{1}{\beta^\alpha \Gamma(\alpha)}\ \mathrm{e}^{\left(\alpha y - e^{y/\beta}\right)} dy$$

To integrate this we use an old trick: we try to leave inside the integral sign the part of the function that integrates to 1 (this is often an easily recognizeable `pdf`); we then pull the constants that get in the way of making this integral equal to 1 out of the integral sign. Luckily, this can be done here:

$$\begin{aligned}
M_Y(t) &= \frac{1}{\beta^\alpha\ \Gamma(\alpha)} \int_{-\infty}^{+\infty} \mathrm{e}^{ty}\ \ \mathrm{e}^{\left(\alpha y - e^{y/\beta}\right)} dy \\
&= \frac{1}{\beta^\alpha\ \Gamma(\alpha)} \int_{-\infty}^{+\infty} \mathrm{e}^{\left(ty + \alpha y - e^{y/\beta}\right)} dy \\
&= \frac{1}{\beta^\alpha\ \Gamma(\alpha)} \cdot \frac{\beta^{\alpha+t}\ \Gamma(\alpha+t)}{\beta^{\alpha+t}\ \Gamma(\alpha+t)} \cdot \int_{-\infty}^{+\infty} \mathrm{e}^{\left(y(\alpha+t) - e^{y/\beta}\right)} dy \\
&= \frac{\beta^{\alpha+t}\ \Gamma(\alpha+t)}{\beta^\alpha\ \Gamma(\alpha)} \cdot \int_{-\infty}^{+\infty} \frac{1}{\beta^{t+\alpha}\ \Gamma(\alpha+t)} \mathrm{e}^{\left(y(\alpha+t) - e^{y/\beta}\right)} dy \\
&= \frac{\beta^t\ \Gamma(\alpha+t)}{\Gamma(\alpha)}
\end{aligned}$$

(a) By definition, the cumulant generating function is simply:

$$S_Y(t) = log(M_Y(t)) = log\left(\frac{\beta^t\,\Gamma(\alpha+t)}{\Gamma(\alpha)}\right)$$

(b) To find the expectation of `Log-Of-Gamma`, we differentiate $S_Y(t) = \log(M_Y(t))$ and evaluate the result at $t = 0$.

$$\begin{aligned}
(S_Y(t))^{'} = (\log(M_Y(t)))^{'} &= \left(\log\left(\frac{\beta^t\,\Gamma(\alpha+t)}{\Gamma(\alpha)}\right)\right)^{'}\\
&= \frac{1}{\frac{\beta^t\,\Gamma(\alpha+t)}{\Gamma(\alpha)}}\cdot\frac{1}{\Gamma(\alpha)}\cdot\left(\beta^t\,\Gamma(\alpha+t)\right)^{'}\\
&= \frac{\Gamma(\alpha)}{\beta^t\,\Gamma(\alpha+t)}\cdot\frac{1}{\Gamma(\alpha)}\cdot\left(\beta^t\log(\beta)\,\Gamma(\alpha+t)+\beta^t\,\Gamma^{'}(\alpha+t)\right)\\
&= \log(\beta)+\frac{\Gamma^{'}(\alpha+t)}{\Gamma(\alpha+t)}
\end{aligned}$$

The last term is the log-derivative of the Gamma function evaluated at $\alpha + t$. This log-derivative is called the `digamma` function and is denoted by $\psi(\cdot)$. Thus:

$$(S_Y(t))^{'} = \log(\beta) + \psi(\alpha + t)$$

We *conclude* that if Y is `Log-Of-Gamma`(α, β) distributed, then its expectation is given by:

$$\boxed{\mathrm{E}\,Y = \log(\beta) + \psi(\alpha)}$$

To find the variance of Y, we we look at the second derivative, $(S_Y(t))^{''}$, and evaluate it at $t = 0$.

$$\begin{aligned}
(S_Y(t))^{''} &= (\log(\beta) + \psi(\alpha + t))^{'}\\
&= \psi^{'}(\alpha + t)
\end{aligned}$$

The derivative of `digamma` is the so-called `trigamma` function:

$$\boxed{Var(Y) = \psi^{'}(\alpha) = \psi_1(\alpha) = \texttt{trigamma}\,(\alpha)}$$

(c) Now we use `limma::trigammaInverse()` to estimate α. Since we estimate $Var(Y)$ by $m_2 - m_1^2$, our estimate of α is:

$$\hat{\alpha} = \texttt{limma::trigammaInverse}(m_2 - m_1^2)$$

And therefore our estimate for β is

$$\log(\hat{\beta}) = m_1 - \texttt{digamma}(\hat{\alpha})$$

or

$$\hat{\beta} = \mathrm{e}^{m_1 - \texttt{digamma}(\hat{\alpha})}$$

(d) Applying the above results to our data, we get:

```
m1 <- mean(gdp)        # = 13478.52
m2 <- mean(gdp^2)      # = 460173780

new.alpha.hat <- limma::trigammaInverse(m2 - m1^2)    # = 5.992181e-05
new.log.of.beta <- m1 - digamma(new.alpha.hat)        # = 30167.52
new.beta.hat <- exp(m1 - digamma(new.alpha.hat))      # Inf
```