Measures of bivariate association

Peter Hoff
STAT 423
Applied Regression and Analysis of Variance
University of Washington
\[ x_c = \text{today's low temp in Minneapolis (in celsius)} \]
\[ y_c = \text{tomorrow's low temp in Green Bay (in celsius)} \]

\text{mean}(x_c \cdot y_c)

## [1] 2.632433
\[ x_k = \text{today's low temp in Minneapolis (in kelvins)} \]
\[ y_k = \text{tomorrow's low temp in Green Bay (in kelvins)} \]

\[
x_k <- x_c + 273.15 ; y_k <- y_c + 273.15
\]
\[
\text{mean}(x_k \cdot y_k)
\]

## [1] 74925.89
\[ \text{mean} \left( (x_k - \text{mean}(x_k)) \times (y_k - \text{mean}(y_k)) \right) \]

## [1] 2.476611

\[ \text{mean} \left( (x_c - \text{mean}(x_c)) \times (y_c - \text{mean}(y_c)) \right) \]

## [1] 2.476611
mean( (xk-mean(xk)) * (yk-mean(yk)) )
## [1] 2.476611

mean( (xc-mean(xc)) * (yc-mean(yc)) )
## [1] 2.476611

cov(xk,yk)
## [1] 2.579803

cov(xc,yc)
## [1] 2.579803

n<-length(xc)
n/(n-1)
## [1] 1.041667

cov(xc,yc) / mean( (xc-mean(xc)) * (yc-mean(yc)) )
## [1] 1.041667
\( x_f = \) today’s low temp in Minneapolis (in fahrenheit)
\( y_f = \) tomorrow’s low temp in Green Bay (in fahrenheit)

\[
\text{xf}\leftarrow xc*9/5+32 \quad \text{yf}\leftarrow yc*9/5+32
\]
\[
\text{mean(xf*yf)}
\]

## [1] 1098.393
\( x_f = \) today’s low temp in Minneapolis (in fahrenheit)

\( y_f = \) tomorrow’s low temp in Green Bay (in fahrenheit)

\[
\text{mean}((x_f - \text{mean}(x_f)) \times (y_f - \text{mean}(y_f)))
\]

```r
## [1] 8.024218
```
\texttt{cor(xc, yc)}

## [1] 0.9138004

\texttt{cor(xk, yk)}

## [1] 0.9138004

\texttt{cor(xf, yf)}

## [1] 0.9138004

\texttt{cor( xc*3.275+1.324 , yc*5.234-6.219)}

## [1] 0.9138004
\texttt{cor(pop,gdp)}

\begin{verbatim}
## [1] 0.5577147
\end{verbatim}

\texttt{cor(log(pop),log(gdp))}

\begin{verbatim}
## [1] 0.803626
\end{verbatim}