Serum vs ABG electrolyte Comparison

Anupam kumar Singh , Imran Khan

26 October 2017

We Aim to analyse Following aspects in this article

1. Distribution of serum and point of care(POC) electrolytes variables
2. Correlation of serum and POC elecyrolytes
3. t. test of serum and POC electrolytes
4. Any misclassification of electrolyte categories by POC vis a vis serum test
5. Deming Regression , Normal regression , Regression equation of serum from POC test
6. Bland -Altman test,graphs, n stats of the two methods

All samples are independent in our article with no repeated measures for any patient

Let us start by creating categories of hypo/hyper/normo for sodium and potassium for both groups

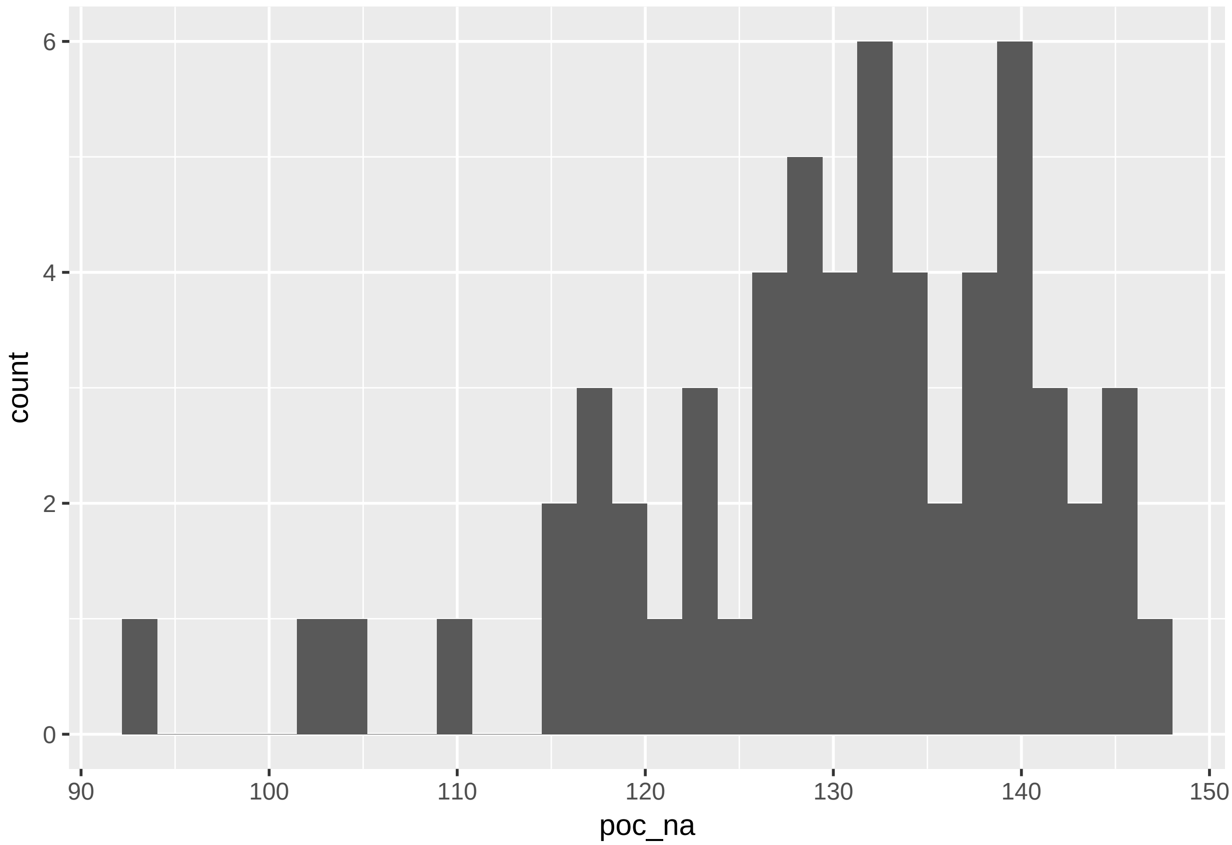
Let us see summary of our data

```
##      poc_na          poc_k          serum_na        serum_k     
##  Min.   : 94.0   Min.   :1.600   Min.   : 96.0   Min.   :2.000  
##  1st Qu.:124.5   1st Qu.:3.400   1st Qu.:123.8   1st Qu.:3.700  
##  Median :132.5   Median :3.950   Median :133.0   Median :4.250  
##  Mean   :130.2   Mean   :4.013   Mean   :130.5   Mean   :4.255  
##  3rd Qu.:138.2   3rd Qu.:4.500   3rd Qu.:138.0   3rd Qu.:4.650  
##  Max.   :148.0   Max.   :6.300   Max.   :148.0   Max.   :6.400  
##      potassium_serum        sodium_serum      potassium_poc
##  Hypokalemia : 9     Hyponatremia :34    Hypokalemia :16  
##  Normal      :42     Normal       :26    Normal      :36  
##  Hyperkalemia: 9     Hypernatremia: 0    Hyperkalemia: 8  
##                                                                
##                                                                
##                                                                
##         sodium_poc
##  Hyponatremia :36  
##  Normal       :24  
##  Hypernatremia: 0  
##                    
##                    
##                    
```

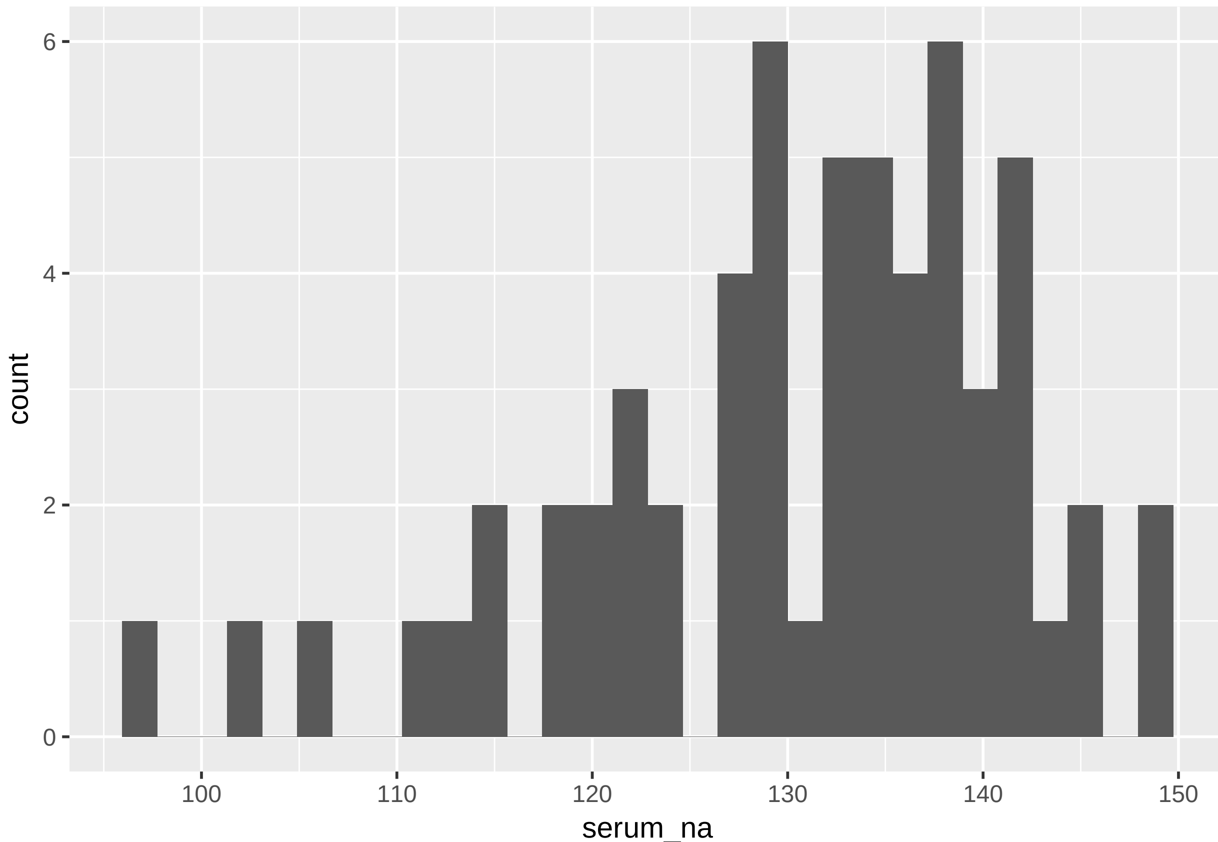
We can see that there are no hpernatremia patients in our population and misclassification is frequent in potassium group.

Let us visualise

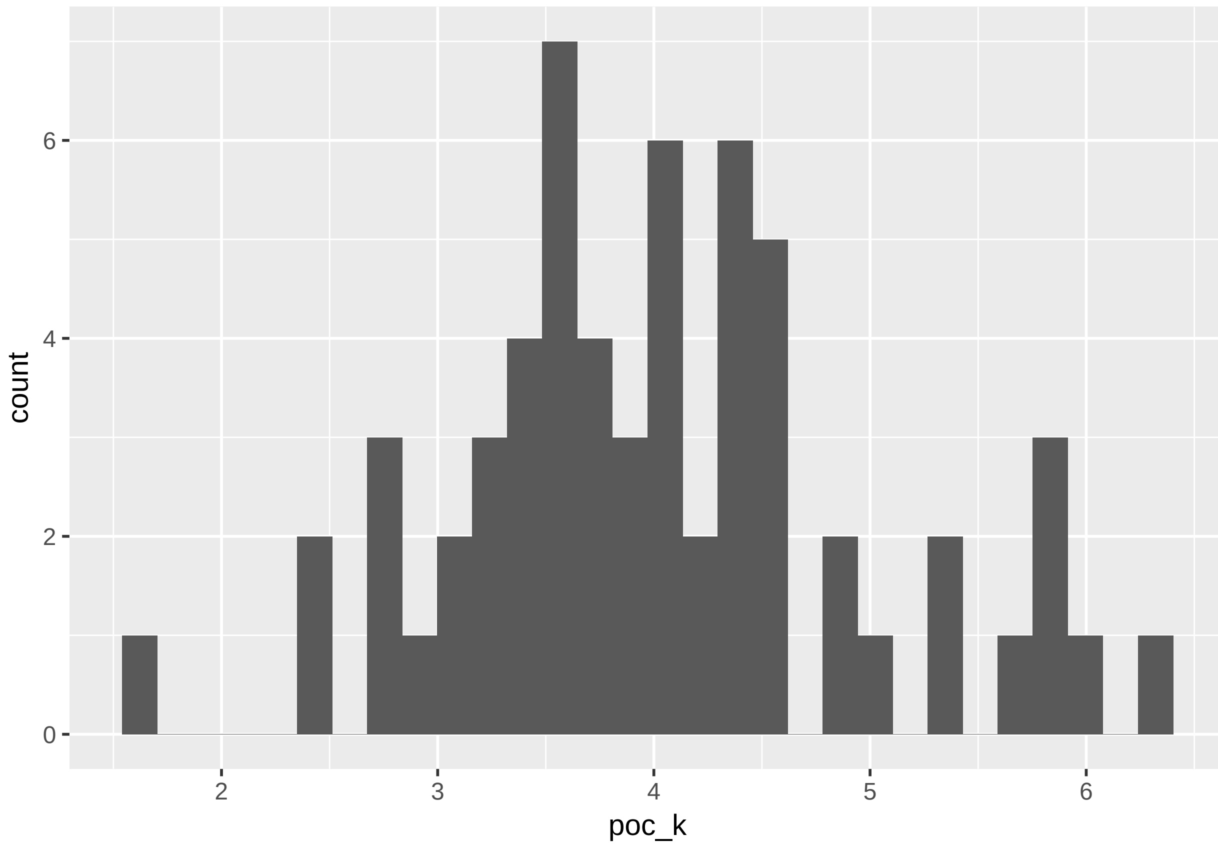
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



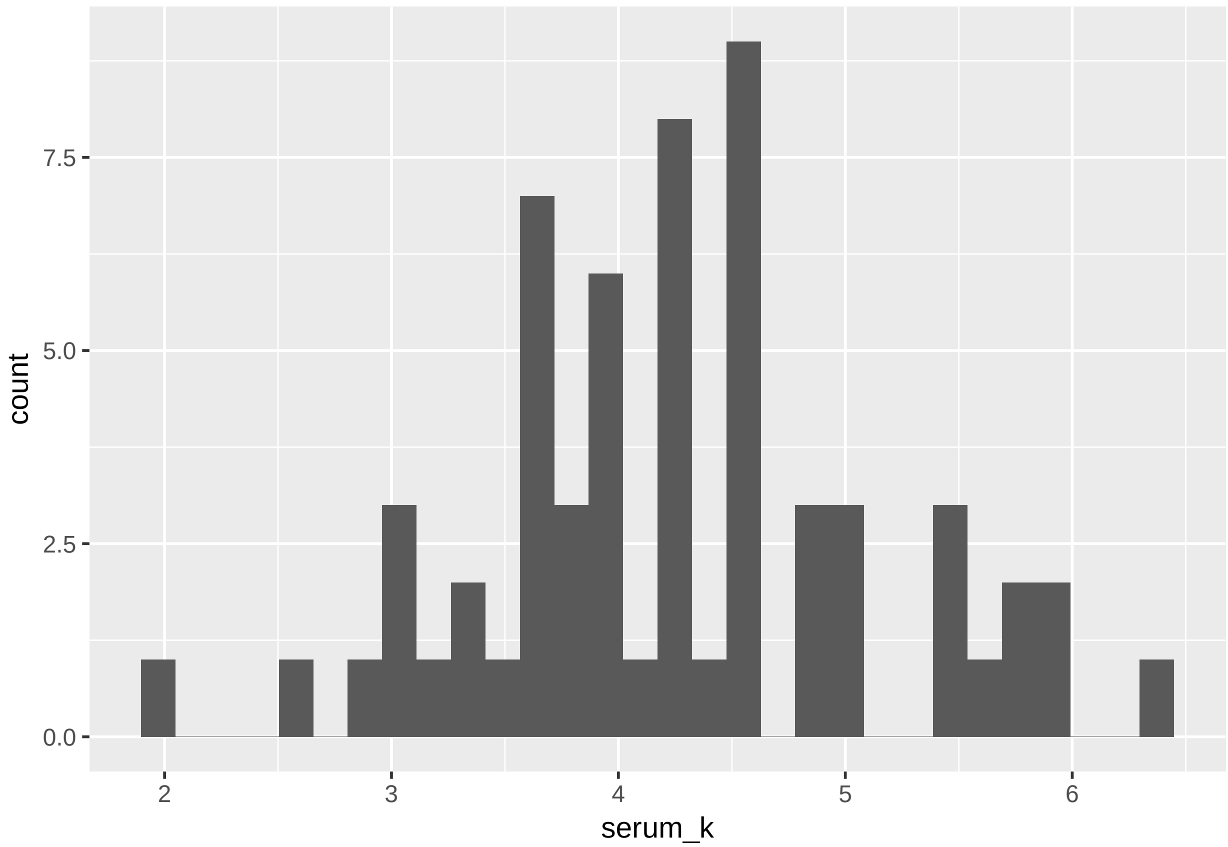

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



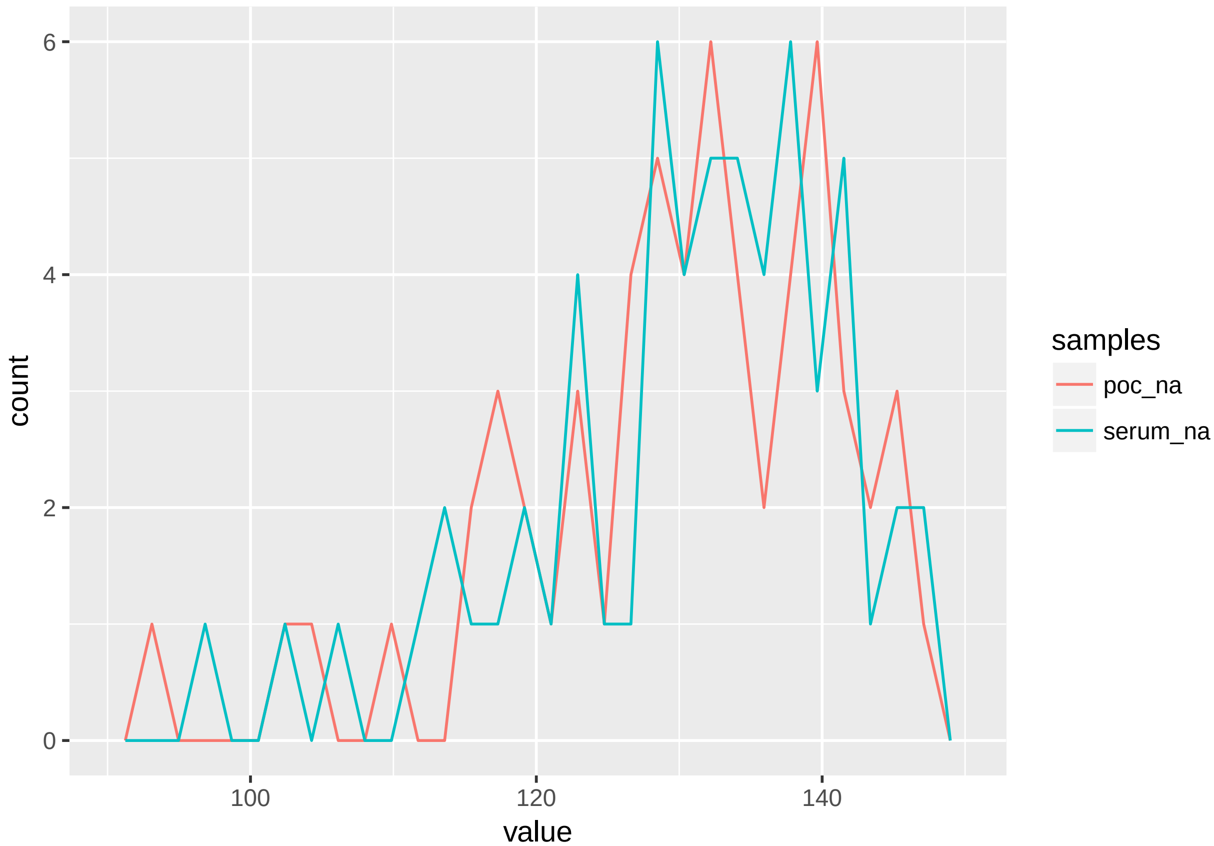

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

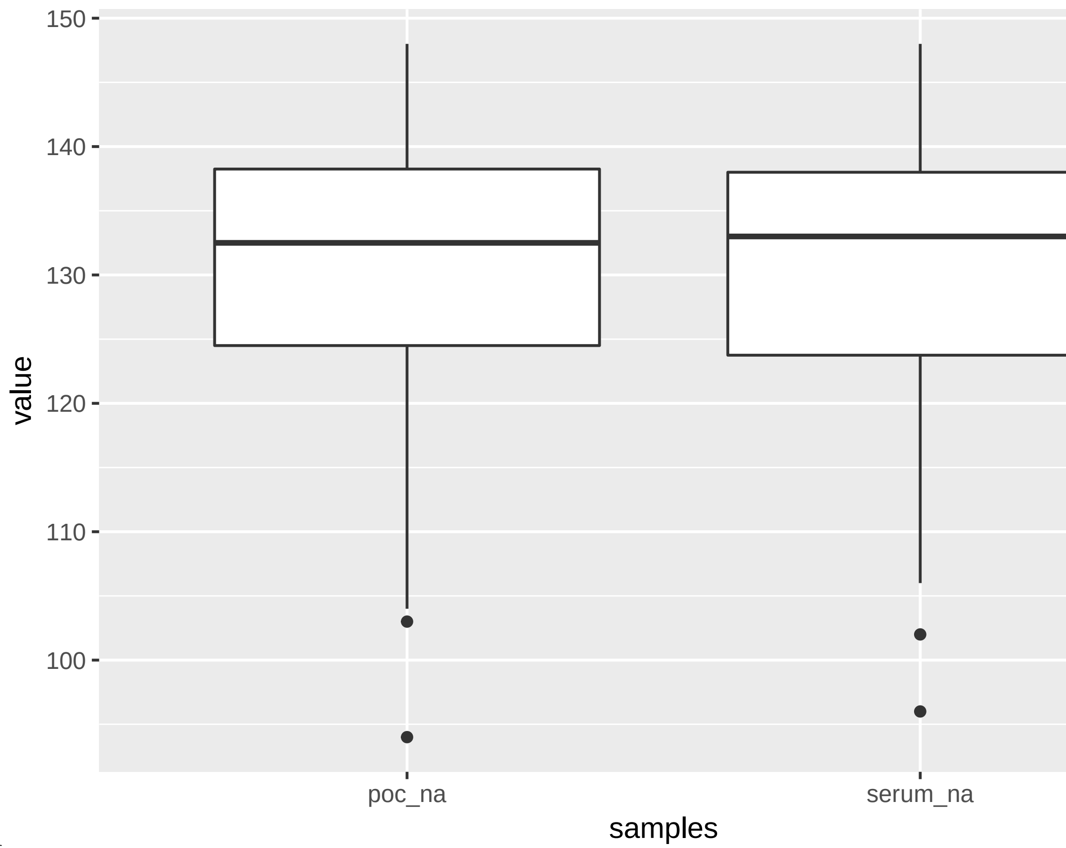


Let us visualise comparative frequency polygons to have relative idea of distribution of sodium

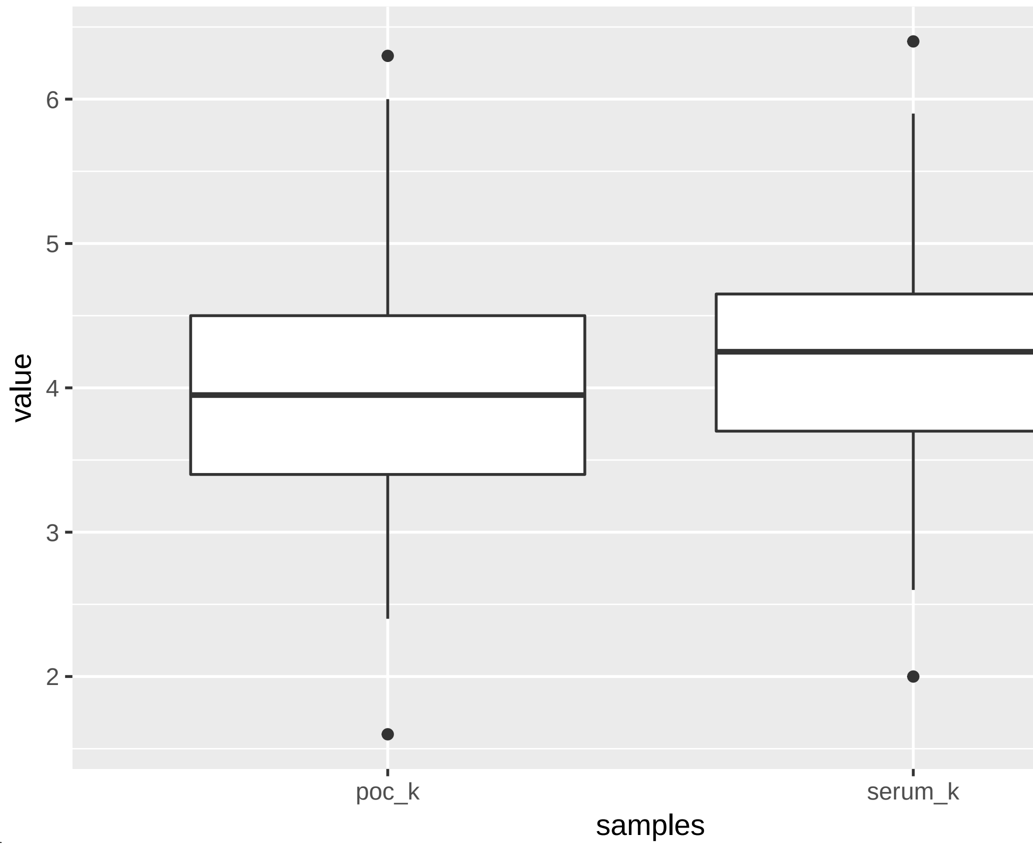
```
electrolytes %>% select(poc_na,serum_na) %>% gather(key="samples",value = "value") %>% ggplot(aes(x=valu
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```





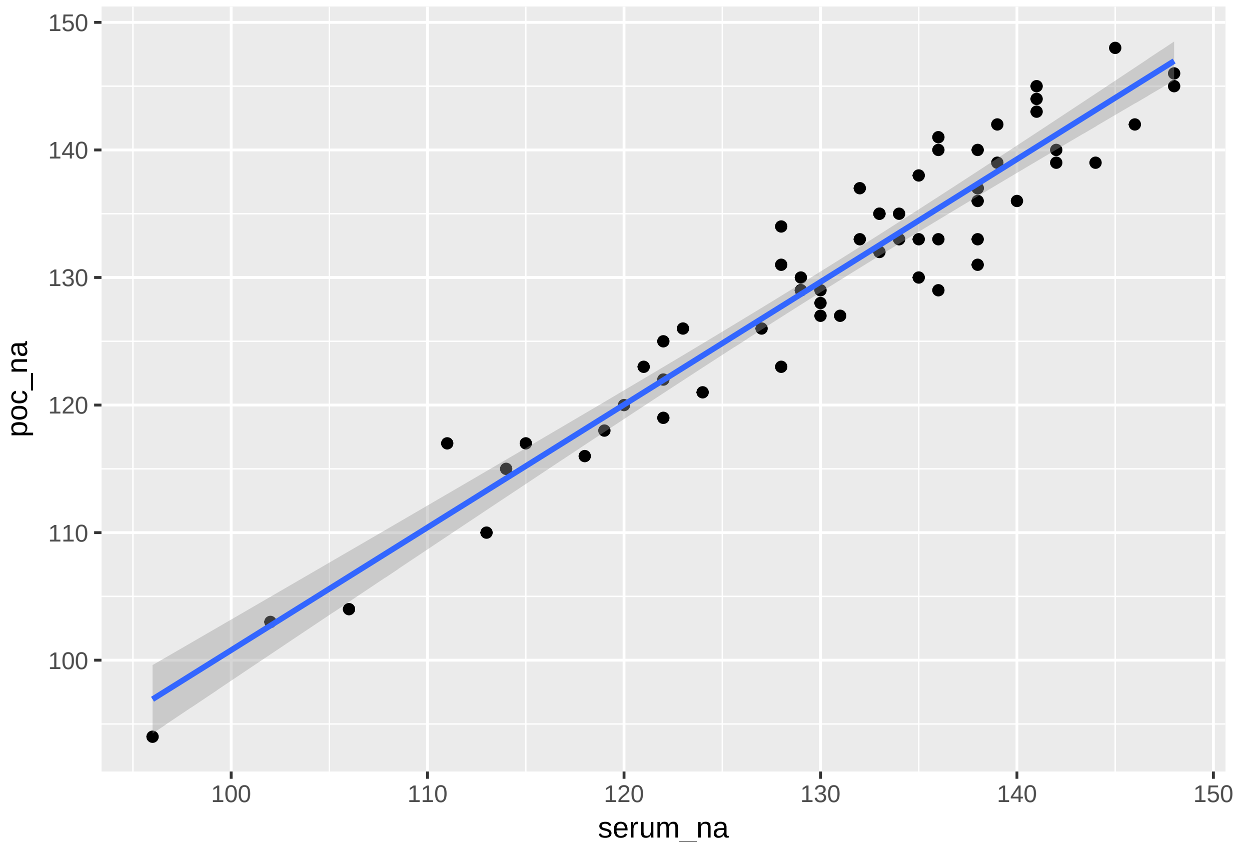
Let us visualise boxplot of sodium



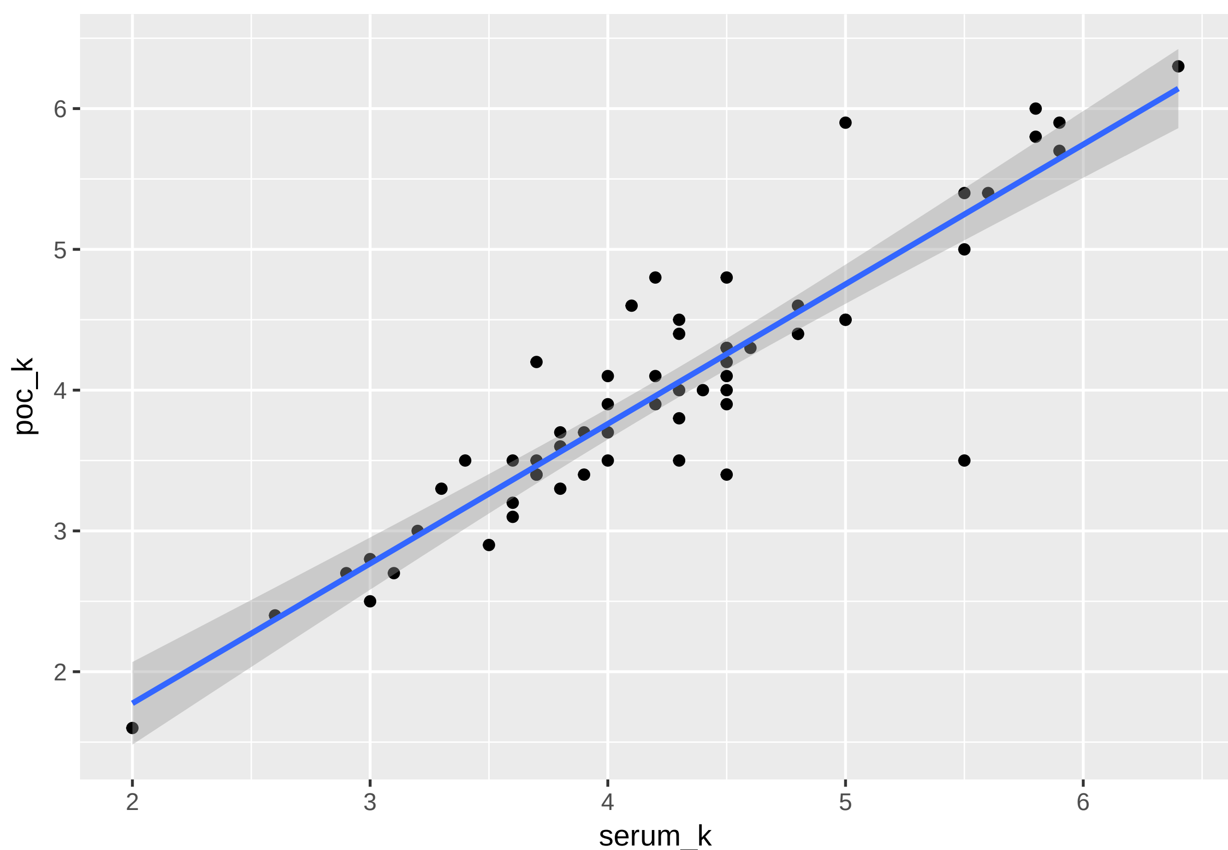
Let us visualise boxplot of potassium

we can see serum potassium is slightly higher than poc potassium.

Now let us visualise scatterplot with regression line for sodium and potassium

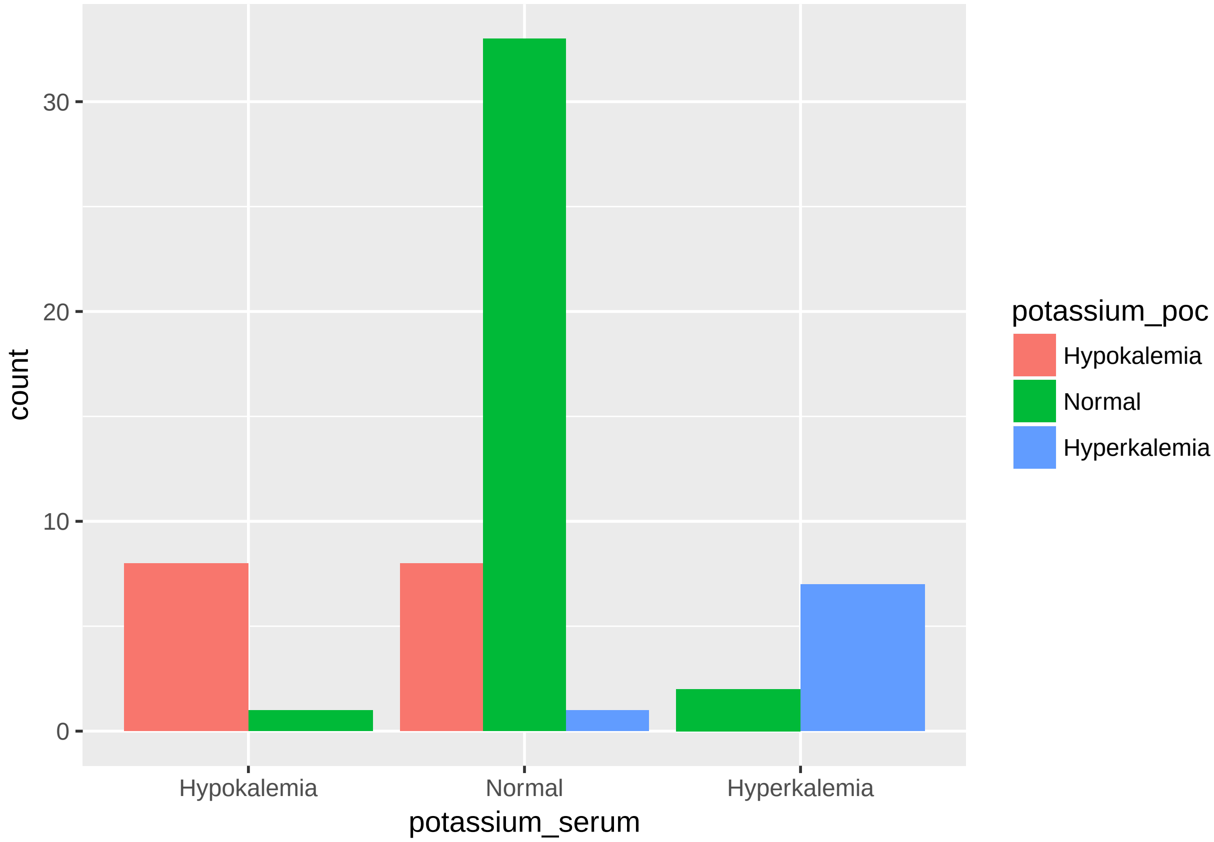


150
140
130
120
110
100
poc_na
100
110
120
130
140
150
serum_na



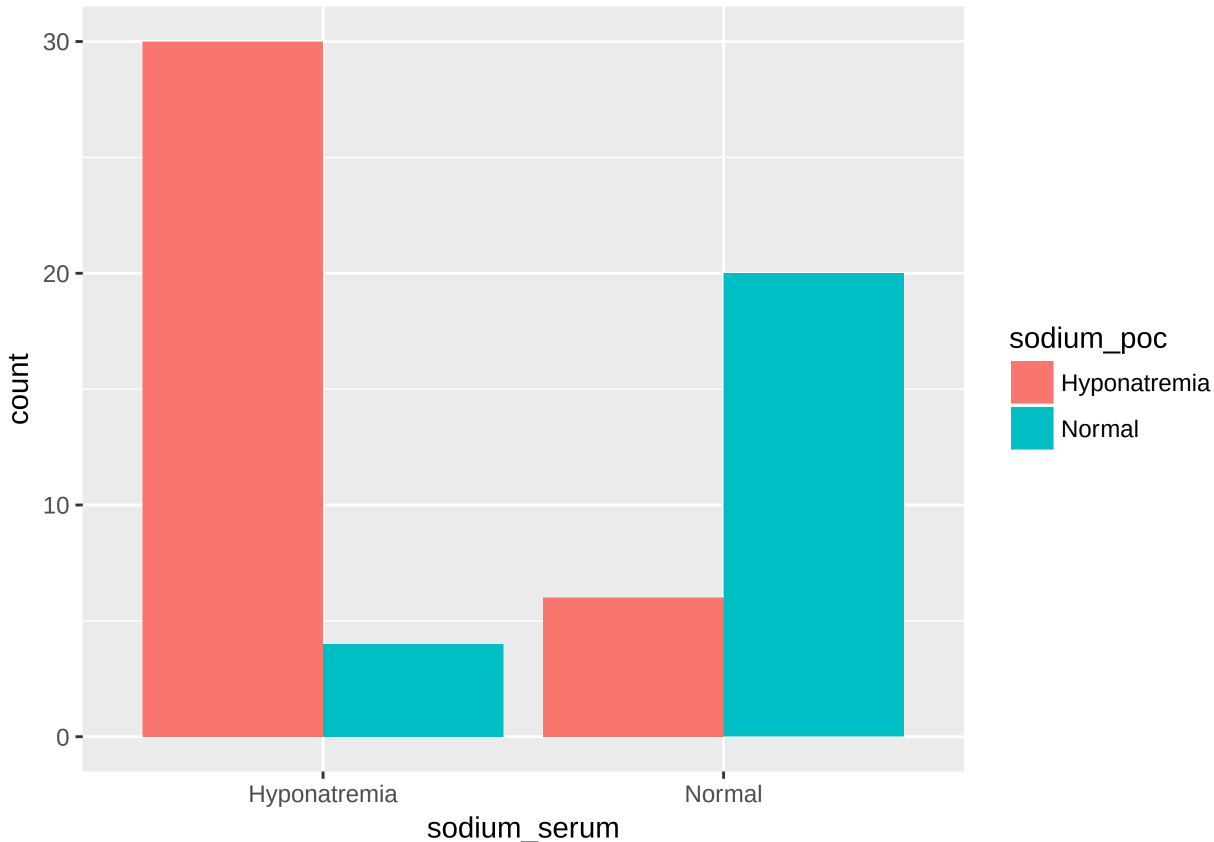
6
5
4
3
2
poc_k
2
3
4
5
6
serum_k

Let us look at classification in potassium

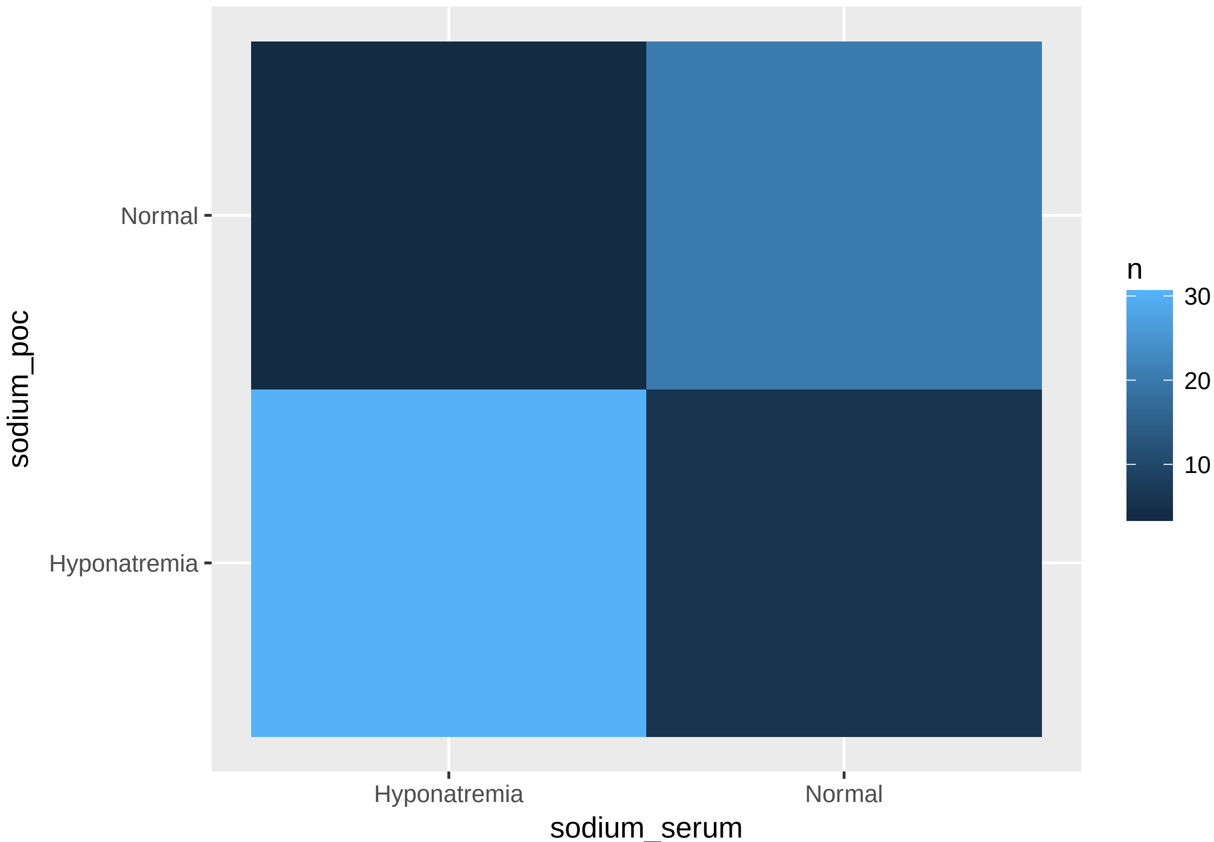


Let us look at classification of sodium

```
electrolytes %>% ggplot(aes(sodium_serum, ..count..)) + geom_bar(aes(fill = sodium_poc), position = "do
```

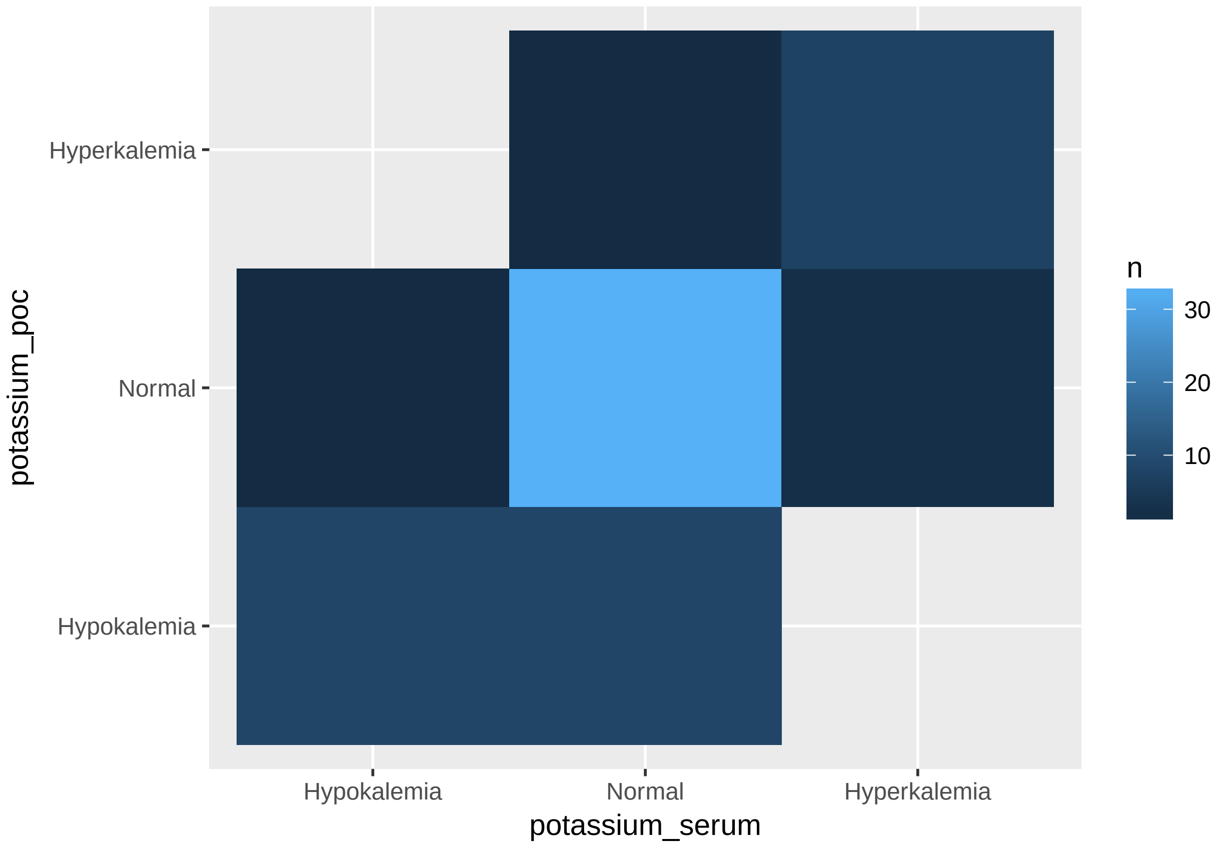


Let us look at heat maps of sodium



Let us look at heat maps of potassium

```
electrolytes %>%
  count(potassium_serum,potassium_poc) %>%
  ggplot(mapping = aes(x = potassium_serum, y = potassium_poc)) +
    geom_tile(mapping = aes(fill = n))
```



Let us look at individual counts

```
electrolytes %>%
  count(sodium_serum,sodium_poc)
```

```
## # A tibble: 4 x 3
##   sodium_serum   sodium_poc     n
##         <fctr>       <fctr> <int>
## 1 Hyponatremia Hyponatremia    30
## 2 Hyponatremia       Normal     4
## 3       Normal Hyponatremia     6
## 4       Normal       Normal    20
```

classification count for potassium

```
electrolytes %>%
  count(potassium_serum,potassium_poc)
```

```
## # A tibble: 7 x 3
##   potassium_serum potassium_poc     n
##            <fctr>        <fctr> <int>
## 1     Hypokalemia   Hypokalemia     8
## 2     Hypokalemia        Normal     1
## 3          Normal   Hypokalemia     8
## 4          Normal        Normal    33
## 5          Normal  Hyperkalemia     1
## 6    Hyperkalemia        Normal     2
## 7    Hyperkalemia  Hyperkalemia     7
```

Now we have visualised data so it is time for some formal statistical tests, First test of skewness

```
electrolytes %>% select(serum_na,serum_k,poc_na,poc_k) %>% map(~skewness(.))
```

```
## $serum_na
## [1] -0.925739
##
## $serum_k
## [1] 0.1768452
##
## $poc_na
## [1] -0.9389183
##
## $poc_k
## [1] 0.3294308
```

We see sodium data is relatively skewed

```
cor.test(electrolytes$poc_na,electrolytes$serum_na,method = "spearman")
```

```
## Warning in cor.test.default(electrolytes$poc_na, electrolytes$serum_na, :
## Cannot compute exact p-value with ties
```

```
##
##  Spearman's rank correlation rho
##
## data:  electrolytes$poc_na and electrolytes$serum_na
## S = 2503.6, p-value < 2.2e-16
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.9304356
```

Next Correlation test for potassium, We have used spearman's rho rank correlation due to non-normal distribution

```
cor.test(electrolytes$poc_k,electrolytes$serum_k)
```

```
##
##  Pearson's product-moment correlation
##
## data:  electrolytes$poc_k and electrolytes$serum_k
## t = 16.367, df = 58, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8479381 0.9433872
## sample estimates:
##       cor
## 0.9066498
```

We can see correlation in potassium is a bit lower than correlation in potassium

Now time for t test for potassium we will use normal paired t test while for sodium wilcox.test due to its non-normal distribution

```
wilcox.test(electrolytes$poc_na,electrolytes$serum_na,paired = TRUE)
```

```
##
##  Wilcoxon signed rank test with continuity correction
```

```
##
## data:  electrolytes$poc_na and electrolytes$serum_na
## V = 672, p-value = 0.4114
## alternative hypothesis: true location shift is not equal to 0
```

we see non-significant difference in sodium

```r
t.test(electrolytes$poc_k,electrolytes$serum_k,paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  electrolytes$poc_k and electrolytes$serum_k
## t = -4.6106, df = 59, p-value = 2.203e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.3465503 -0.1367830
## sample estimates:
## mean of the differences
##              -0.2416667
```

we see that serum potassium is signifcantly higher than poc potassium, notably it was obvious from boxplot as well ..now we have a statistical test to say the same

Now let us look at regression equation for serum sodium and potassium to see if they can be predicted from poc test

First for potassium

```r
f = lm(serum_k~poc_k,data=electrolytes)
summary(f)
```

```
##
## Call:
## lm(formula = serum_k ~ poc_k, data = electrolytes)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8176 -0.2193  0.0131  0.1709  1.6702
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.93101    0.20876    4.46 3.82e-05 ***
## poc_k        0.82824    0.05061   16.37  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.374 on 58 degrees of freedom
## Multiple R-squared:  0.822,  Adjusted R-squared:  0.8189
## F-statistic: 267.9 on 1 and 58 DF,  p-value: < 2.2e-16
```

Next for sodium

```r
fna = lm(serum_na~poc_na,data=electrolytes)
summary(fna)
```

```
##
## Call:
## lm(formula = serum_na ~ poc_na, data = electrolytes)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9240 -2.5405  0.1098  2.1820  6.6694
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.88257    4.79711   1.226    0.225
## poc_na       0.95762    0.03673  26.075   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.143 on 58 degrees of freedom
## Multiple R-squared:  0.9214, Adjusted R-squared:   0.92
## F-statistic: 679.9 on 1 and 58 DF,  p-value: < 2.2e-16
```

we see adjusted R square is 0.92 for sodium and 0.81 for potassium implying POC sodium has higher predictive value for serum compared to potassium

We would also like to use deming regression for this analysis as it accounts for error in both variables and is required in clinical chemistry. you can find details here and here

First Deming regression for sodium

```r
dem.sodium <- mcreg(electrolytes$serum_na,electrolytes$poc_na,method.reg = "Deming")
dem.sodium@para
```

```
##                  EST SE        LCI      UCI
## Intercept -0.6897265 NA -9.1897031 8.765223
## Slope      1.0024752 NA  0.9305773 1.067353
```

Next Deming regression for Potassium

```r
dem.potassium <- mcreg(electrolytes$serum_k,electrolytes$poc_k,method.reg = "Deming")
dem.potassium@para
```

```
##                  EST SE       LCI        UCI
## Intercept -0.6879562 NA -1.167004 -0.3418846
## Slope      1.1048859 NA  1.006981  1.2258538
```

Deming regressions are of historical importance, but important because sometimes journals can ask for them

Now let us look at classification accuracy by kappa measurement. You can read more about it here

Cohen's kappa for Sodium

```r
elec_sod =electrolytes %>% select(sodium_poc,sodium_serum)
cohen.kappa(elec_sod)
```

```
## Warning in any(abs(bounds)): coercing argument of type 'double' to logical
```

```
## Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels)
##
## Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
##                  lower estimate upper
## unweighted kappa  0.46     0.66  0.85
## weighted kappa    0.46     0.66  0.85
##
##  Number of subjects = 60
```

Cohen's kappa for Potassium

```r
elec_pot =electrolytes %>% select(potassium_poc,potassium_serum)
cohen.kappa(elec_pot)
```

```
## Warning in any(abs(bounds)): coercing argument of type 'double' to logical

## Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels)
##
## Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
##                  lower estimate upper
## unweighted kappa  0.42     0.62  0.81
## weighted kappa    0.52     0.67  0.83
##
##  Number of subjects = 60
```

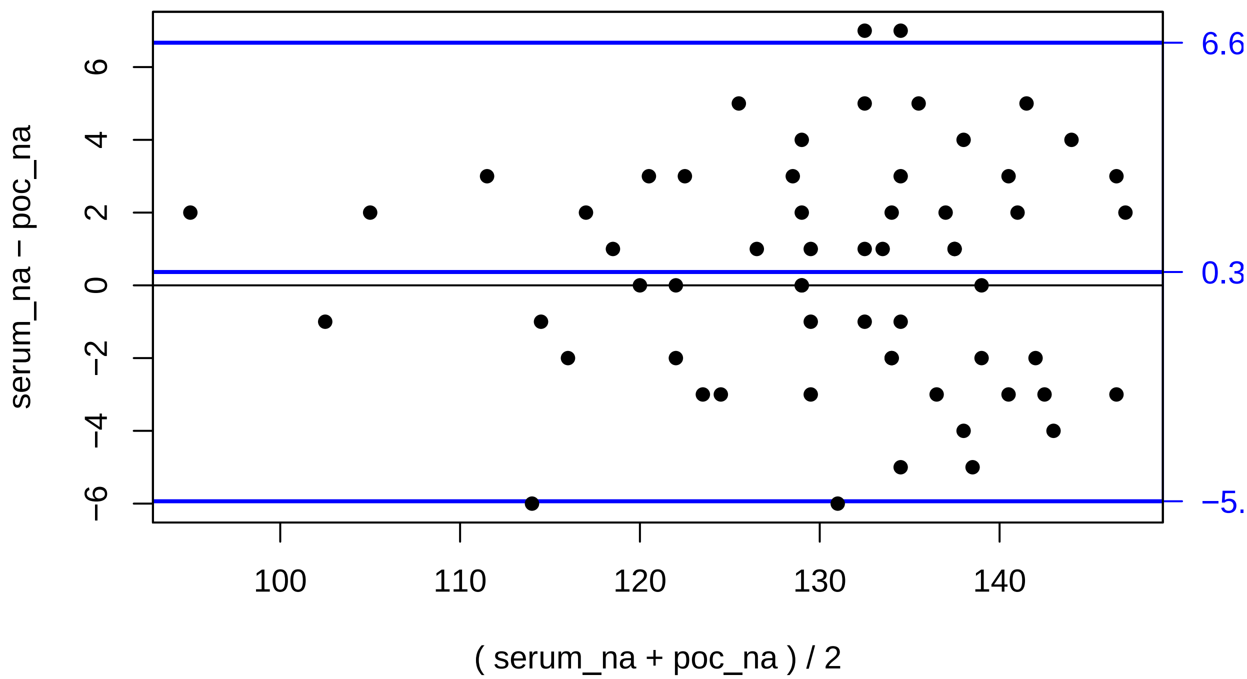
As expected cohen's kappa is slightly lower for potassium than sodium in conconcordance with our observations till now.

Now we will calculate Bland Altman Plot and statistics for sodium and potassium . You can read more about it in this article and this article

Bland Altman Plot of Sodium

```r
 with(electrolytes,BlandAltman(serum_na,poc_na))
```

```
## NOTE:
##  'AB.plot' and 'BlandAltman' are deprecated,
##  and likely to disappear in a not too distant future,
##  use 'BA.plot' instead.
```



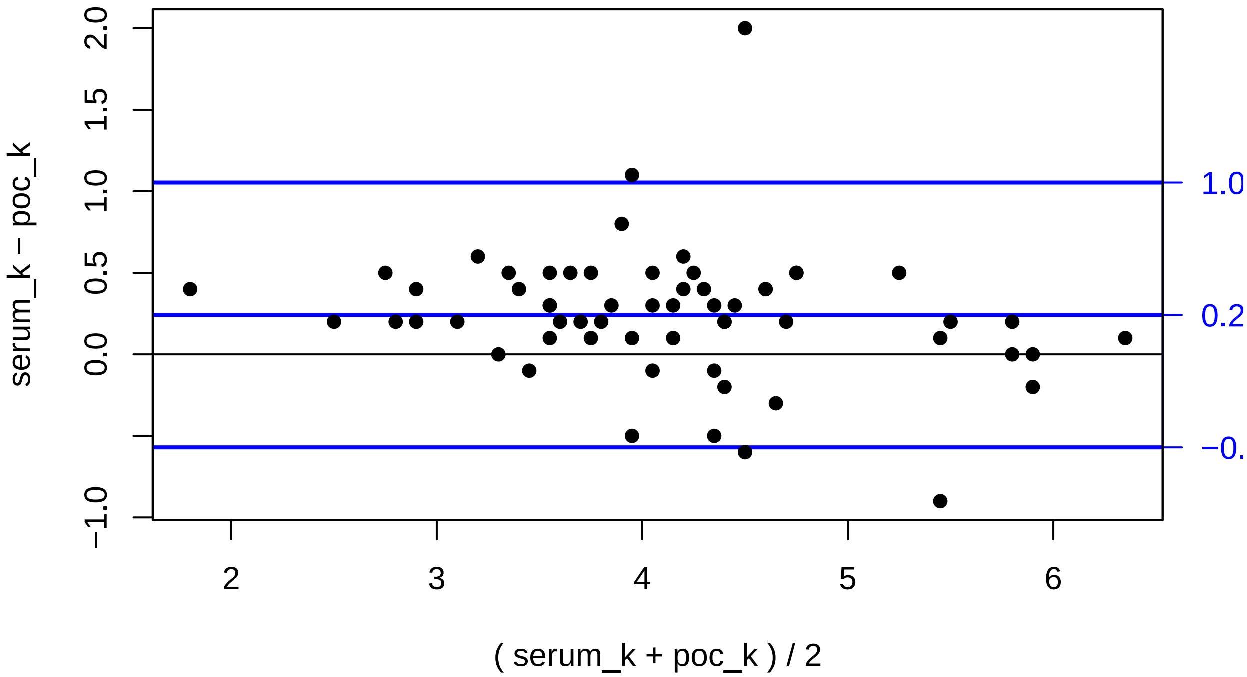

```
##
## Limits of agreement:
## serum_na - poc_na        2.5% limit       97.5% limit          SD(diff)
##         0.3666667        -5.9360545         6.6693878         3.1513606
```

Thus For sodium Serum sodium is higher than POC test by 0.36 and 95% C.I is likely to be between - 5.93 to 6.66 , It shoud be kept in mind that limit of allowable bias for sodium is 4 meq/L

Now Bland Altman Plot of Potassium

```
with(electrolytes,BlandAltman(serum_k,poc_k))
```

```
## NOTE:
##  'AB.plot' and 'BlandAltman' are deprecated,
##  and likely to disappear in a not too distant future,
##  use 'BA.plot' instead.
```




```
##
## Limits of agreement:
## serum_k - poc_k      2.5% limit     97.5% limit        SD(diff)
##       0.2416667      -0.5703546       1.0536879       0.4060106
```

Thus For potassium Serum potassium is higher than POC test by 0.24 and 95% C.I is likely to be between -0.57 to 1.0 , It shoud be kept in mind that limit of allowable bias for sodium is 0.5 meq/L