

IntermediateProblemAnswers

May 12, 2018

1 Intermediate Problem Answers

1.1 Regression Problem

```
In [ ]: ##### Prepare Data

X = rand(1000, 3)           # feature matrix
a0 = rand(3)                # ground truths
y = X * a0 + 0.1 * randn(1000); # generate response

X2 = hcat(X, ones(1000))
println(X2\y)

using MultivariateStats
println(llsq(X,y))

using DataFrames, GLM
data = DataFrame(X1=X[:,1], X2=X[:,2], X3=X[:,3], Y=y)
OLS = lm(@formula(Y ~ X1 + X2 + X3), data)

X = rand(100);
y = 2X + 0.1 * randn(100);

using Plots
b = X\y
println(b)
gr()
scatter(X,y)
Plots.abline!(b[1],0.0, lw=3) # Slope, Intercept
```

1.2 Distribution Dispatch Problem

This is from Josh Day's talk: <https://www.youtube.com/watch?v=EwcTNzpQ6Sc>

Solution is from: https://github.com/joshday/Talks/blob/master/SLG2016_IntroToJulia/Slides.ipynb

```
In [ ]: function myquantile(d::UnivariateDistribution, q::Number)
    θ = mean(d)
```

```

    tol = Inf
    while tol > 1e-5
         $\theta$ old =  $\theta$ 
         $\theta$  =  $\theta$  - (cdf(d,  $\theta$ ) - q) / pdf(d,  $\theta$ )
        tol = abs( $\theta$ old -  $\theta$ )
    end
     $\theta$ 
end

for dist in [Gamma(5, 1), Normal(0, 1), Beta(2, 4)]
    @show myquantile(dist, .75)
    @show quantile(dist, .75)
    println()
end

```