Math130 L113 4/10/12

Goodness of fit

actual data fits on assumed distribution

equal frequency, uniformly distributed $E = \frac{n}{k}$

% for each category $E = np$

T.S. $\chi^2 = \sum \frac{(O-E)^2}{E}$

Tire	LF	RF	LR	RR	$E = \frac{40}{4} = 10$
O	11	15	8	6	
① E	10	10	10	10	

Claim = uniform distribution $E = \frac{n}{k}$

② TS $\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(11-10)^2}{10} + \frac{(15-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(6-10)^2}{10}$

$= 4.6$

③ P-value 2nd vars ⇧8: χ^2cdf(TS, 999, df)

$df = k-1$ (4.6, 999, 3)

④ $p = 0.204$

p-value ≤ α yes not a good fit

no is a good fit

$0.204 \leq 0.05$ no, is a good fit

distribution is uniform

Digit	1	2	3	4	5	6	7	8	9	
O	0	15	0	76	476	183	8	23	0	$n = 784$
① E	236	138	98	76	62	53	45	40	36	$E = np$

	1	2	3	4	5	6	7	8	9
%	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6

$784(.301) \approx 236$
$784(.176) \approx 138$
$784(.125) \approx 98$
$784(.097) \approx 76$
$784(.079) \approx 62$
$784(.067) \approx 53$
$784(.058) \approx 45$
$784(.051) \approx 40$
$784(.046) \approx$

② TS: $\chi^2 = \sum \frac{(O-E)^2}{E} = 3600.6$

③ p-value: $\chi^2 cdf(3600.6, 9999, 8) = 0$ (df $= 9-1$)

④ p-value $\leq \alpha$ $0 \leq 0.01$ yes not a good fit
not same distribution
checks result of fraud

	Basketball	Baseball	Hockey	Football
Home	127	53	50	57
Visitor	91	47	43	42

① TS: $\chi^2 = 4.737$

② P-value: $p = 0.192$

③ P-value $\leq \alpha$ $0.192 \leq .05$ no, independent

④ Expected values

115.97	58.571	54.471	57.986
82.029	41.429	38.529	41.014

home/visitor wins independent of sport