THE POISSON PROCESS

LEARNING OBJECTIVES STOCHASTIC PROCESSES	KNOWLEDGE STATEMENTS
1. Describe the properties of Poisson processes: • For increments in the homogeneous case • For interval times in the homogeneous case • For increments in the non-homogeneous case • Resulting from special types of events in the Poisson process • Resulting from sums of independent Poisson processes Range of weight: 0-5 percent	a. Poisson process b. Non-homogeneous Poisson
2. For any Poisson process and the inter-arrival and waiting distributions associated with the Poisson process, calculate: • Expected values • Variances • Probabilities Range of weight: 0-5 percent	a. Probability calculations for Poisson process
3. For a compound Poisson process, calculate moments associated with the value of the process at a given time Range of weight: 0-5 percent	a. Compound Poisson process
READINGS: Daniel Poisson	

©ACTEX 2014

1. THE POISSON DISTRIBUTION

The parameter λ can be thought of as a rate at which events occur in a time interval, and the random variable X represents the number of events which occur in the time interval.

$$P(X=k)=p(k)=\left(\frac{e^{-\lambda}\lambda^{k}}{k!}\right) \tag{1.1}$$

$$\left[E(X)=V(X)=\lambda\right] \tag{1.2}$$

$$M(t)=e^{\lambda(e^{t}-1)} \tag{1.3}$$

Poisson Approximation to the Binomial

For large n it can be difficult to calculate values of $\binom{n}{k}$ needed for binomial calculations. There is a cut-off point where n gets too large for your calculator to handle exactly. For example, the TI BA II Plus Professional returns an error message for the calculation of $\binom{500}{250}$.

Thus the following result is helpful:

For large n and small p, binomial probabilities can be approximated as Poisson Probabilities with $\lambda = np$. (1.4)

Example 1-1 Let $n = 100$ and $p = 0.01$. Then $\lambda = np = 1$. We compare $p(0)$, $p(1)$, and $p(2)$ calculations using the TI BAII Plus Professional:

Binomial; $n = 100$, $p = 0.01$	Poisson; $\lambda = 1$
$p(0)=\binom{100}{0}(0.99)^{100}(0.01)^{0}=0.3660$	$p(0)=\frac{e^{-1}(1)^{0}}{0!}=0.3679$
$p(1)=\binom{100}{1}(0.99)^{99}(0.01)^{1}=0.3697$	$p(1)=\frac{e^{-1}(1)^{1}}{1!}=0.3679$
$p(2)=\binom{100}{2}(0.99)^{98}(0.01)^{2}=0.1849$	$p(2)=\frac{e^{-1}(1)^{2}}{2!}=0.1839$

Further discussion can be found in reference Probability for Risk Management, Section 5.3.2.

©ACTEX 2014

Exercise 1

If $n = 200$ and $p = 0.008,$ use the Poisson probability to approximate the binomial probability $p(2)$.

Answer: 0.25843

Sums of Independent, Poisson Random Variables

From our study of probability, we know that if $N_1, N_2, \ldots, N_k$ are independent Poisson random variables with $E(N_i) = \lambda_i$

$$\begin{aligned} N &= N_1 + \cdots + N_k \text{ is Poisson with parameter} \\ \lambda &= E(N) = \lambda_1 + \cdots + \lambda_n \end{aligned} \tag{1.5}$$

That is, the sum of Poisson random variables is also Poisson.

Example 1-2

The numbers of male and female customers arriving at a bank between 10 AM and 11 AM are independent Poisson random variables with $E(N_2) = 50.$ and $E(N_2) = 50.$ Then the total number of arrivals

$$N = N_1 + N_2$$

is a Poisson random variable with $E(N) = 50 + 50 = 100$.

Decomposition of Poisson Random Variables

Example 1-3

We now look at the bank in the previous example in a different way. Suppose an analyst provides the information that:

a) the total number of arrivals is Poisson with $\lambda = 100$, and
b) 50% of the arrivals are male and 50% are female.

Here, the random variable N is given, but the male and female arrival random variables N_1 and N_2 are not given. The most intuitive step would be to guess that N_1 is Poisson with $\lambda_1 = 0.50(100) = 50$ and N_2 is Poisson with $\lambda_2 = 0.50(100) = 50$. The following decomposition theorem shows that if male and female arrivals are independent, this guess is correct.

Example 1-4

Let N be a Poisson random variable with parameter λ. Suppose that each event can be classified as one and only one of k different types independently of all other events. Denote the probability that an event is type i by p_i, and N_i denote the random variable for events of type i.

©ACTEX 2014

Then, $N_1, \ldots, N_k$ are independent Poisson distributions with $E(N_i) = \lambda_i = \lambda\, p_i$.

Example 1-5

The random variable N for accident claims at a company is Poisson with $\lambda = 100$ per week. The insured population is classified into three categories: preferred (P), standard (S), and high-risk (HR). 10% of the claims come from preferred insureds, 30% from standard insureds and the remaining 60% from high-risk insureds.

Note that we are not stating that 10% of the insureds are preferred, but only that the preferred insureds have 10% of the claims.

Let N_1, N_2, and N_3 be the independent claim frequency random variables for P, S, and HR, respectively. Then, N_1, N_2, and N_3 are Poisson random variables with parameters $\lambda_1 = 10$, $\lambda_2 = 30$, and $\lambda_3 = 60$. Total Claims: Poisson N, $\lambda = 100$

The breakdown into components is:

Preferred	Standard	High Risk
$p_1 = 0.10, \lambda_1 = \lambda p_1 = 10$	$p_2 = 0.30, \lambda_2 = \lambda p_2 = 30$	$p_3 = 0.60, \lambda_3 = \lambda p_3 = 60$
$N_1, \lambda_1 = 10$	$N_2, \lambda_2 = 30$	$N_3, \lambda_3 = 60$

Exercise 2

Vehicles arrive at an intersection at a Poisson rate of 60 per hour. 70% of these are cars and 30% are trucks. What are the frequency random variables N_c and N_t for the arrivals of cars and trucks?

Answer: N_c : Poisson, $\lambda_c = 42$, N_t : Poisson, $\lambda_t = 18$

We will show in a later section that the Poisson process with rate λ per unit time interval can be converted to a Poisson process with rate λt per t time intervals under reasonable assumptions. In the example we have just seen, we might convert $\lambda = 100$ per business week to $\lambda = 100/5$ per business day. The next exercise uses this decomposition.

Example 1-6
Fall 2003 #11 SOA exam

Subway trains arrive at a station at a Poisson rate of 20 per hour. 25% of the trains are express and 75% are local. The type of each train is independent to the types of preceding trains.

An express gets you to the stop for work in 16 minutes and a local gets you there in 28 minutes. You always take the first train to arrive, and your co-worker takes the first express. You both are waiting at the same station. Calculate the probability that the train you take will arrive at the stop before the train your co-worker takes.

Solution

If an express arrives first you will both get on it, and you will not arrive first. The only way for you to arrive first is for both of the following events to occur:

©ACTEX 2014

A) A local arrives first.
B) No express arrives in the next 12 minutes (since 28–16 = 12).

$P(A)$ = P(a local arrives first) = 0.75 . Let N_1 and N_2 be the arrival number random variables for the expresses and locals respectively. **Now, we use the decomposition**:

(express) N_1 is Poisson with $\lambda_1 = (0.25)(20) = 5$

(local) N_2 is Poisson with $\lambda_2 = (0.75)(20) = 15$

$P(A) = P$(a local arrives first) $= 0.75$

To find $P(B)$, we adjust our Poisson rate per hour to a rate per 12 minutes, which is $1/5$ of an hour. $\lambda_1 = 5$ per hour, so for a 12 minute period: $\lambda_{12} = 5/5 = 1$ per 12-minute period. The express arrivals for a 12-minute period, are Poisson X_{12} with parameter $\lambda_{12} = 1$.

$$P(B) = P(\text{no express in a 12 min. period}) = P(X_{12} = 0) = \frac{e^{-1}1^0}{0!} = 0.3679$$

Since A and B are independent, our final answer is:

$$P(A \,\&\, B) = P(A)P(B) = 0.75(0.3679) = 0.2759$$

Note: In the previous example, we calculated P(B) using Poisson probabilities. Alternatively, we could use the relation between the Poisson and exponential distributions as reviewed in Module 2. Here is the alternate approach:
If a Poisson process has rate λ, waiting times between events are exponentially distributed with $\theta = 1/\lambda$, and $F(x) = 1 - e^{-(x/\theta)}$. Thus with $\lambda = 5$ for an express, waiting time X is exponential with $\theta = 1/5$.

$$P(B) = P\left(\text{waiting time} > \tfrac{1}{5}\text{hour}\right) = 1 - F\left(\tfrac{1}{5}\right) = S\left(\tfrac{1}{5}\right) = e^{\frac{-(1/5)}{(1/5)}} = e^{-1}, \text{ as previously calculated.}$$

Exercise 3

For the vehicles arriving at the intersection in ***Exercise 2***, what is the probability that in the next 6 minutes exactly 2 cars and no trucks arrive?

Answer: 0.02186

©ACTEX 2014

2. CREATING NEW COUNTING MODELS

Compound Frequency Models

Consider compound sums of the form

$$S = X_1 + \cdots + X_N,$$

where the random variables $N, X_1, \ldots, X_N$ are independent. It is typically the case that N, the primary random variable, gives a claim count and the secondary random variables $X_1, \ldots, X_N$ represent claim amounts. However, it is possible for the X_i to be themselves count distributions – in which case the compound random variable S would be a compound count distribution. We will give an example of a compound distribution here, in order to review the double expectation theorem and the normal approximation.

Example 2-1
Fall 2001 #30 SOA /CAS

The claims department of an insurance company receives envelopes with claims for insurance coverage at a Poisson rate of $\lambda = 50$ envelopes per week. For any period of time, the number of envelopes and the number claims in the envelopes are independent. The numbers of claims in the envelopes have the following distribution:

Number of Claims	Probability
1	0.20
2	0.25
3	0.40
4	0.15

Using the normal approximation, calculate the 90th percentile of the number of claims received in 13 weeks.

Solution:
Let N be the number of envelopes and X_i the number of claims in the i^{th} envelope. Then the total number of claims is

$$S = X_1 + \cdots + X_N.$$

The X_i are *iid* with distribution given in the preceding table, and the X_i are independent of N. Since N is Poisson, S has a compound Poisson distribution.

Note that such problems must be read with care. The Poisson rate of $\lambda = 50$ given in the problem is for one week. The rate for a 13-week period is

$$\lambda = 13(50) = 650$$

©ACTEX 2014

To use the normal approximation, we must find the mean and standard deviation of *S*. These can be found using the double expectation theorem, given by

$$E(S) = \lambda E(X)$$
$$V(S) = \lambda E(X^2)$$

From the normal table, $E(X) = 1(0.20) + 2(0.25) + 3(0.40) + 4(0.15) = 2.5$

$$E(X^2) = 1(0.20) + 4(0.25) + 9(0.40) + 16(0.15) = 7.2$$

Thus $E(S) = 650(2.5) = 1,625$

$$V(S) = 650(7.2) = 4,680$$
$$\sigma_S = \sqrt{4,680} = 68.4$$

The 90^{th} percentile of the standard normal is $z_{0.90} = 1.282$.

Thus the 90^{th} percentile for *S* is $S_{0.90} = \mu_S + z_{0.90}\sigma_S$

$$= 1,625 + 1.282\left(\sqrt{4,680}\right)$$
$$= 1,712.70$$

Exercise 4

A town sells lottery tickets. A winning ticket pays 1 with probability 0.70, 5 with probability 0.2, and 10 with probability 0.10. The number of winning tickets has a Poisson distribution with rate 50 per month. Using the normal approximation, find the 95th percentile of total payout in a year.

Answer: 1779.66

Example 2.1 used the normal approximation for a large λ of 650. One could also calculate exact probabilities for *S* using an intuitive approach. For example,

i) The only way to have no claims is to have no envelopes: $P(S = 0) = P(N = 0)$

ii) The only way to have exactly one claim is to have exactly one envelope with exactly one claim: $P(S = 1) = P(N = 1)P(X_1 = 1 \mid N = 1) = (650)e^{-650}(0.20)$



Important Note!

It can be tedious to generate probabilities for compound counting distributions as we have just done. However for smaller values of λ the normal approximation does not apply, and we must generate probabilities directly.

©ACTEX 2014

PROBLEMS SET 1

1. Spring 2000 #2 SOA

Lucky Tom finds coins on his way to work at a Poisson rate of 0.5 coins/minute. The denominations are randomly distributed:

(i) 60% of the coins are worth 1;

(ii) 20% of the coins are worth 5; and

(iii) 20% of the coins are worth 10

Calculate the conditional expected value of the coins Tom found during his one-hour walk today, given that among the coins he found exactly ten were worth 5 each.

(A) 108

(B) 115

(C) 128

(D) 165

(E) 180

2. Fall 2003 #12 CAS

A driver is selected at random. If the driver is a "good" driver, he is from a Poisson population with a mean of 1 claim a year. If the driver is a "bad" driver, he is from a Poisson population with a mean of 5 claims per year. There is an equal probability that the driver is either a "good" or a "bad" driver. If the driver had 3 claims last year, calculate the probability that the driver is a "good" driver.

(A) Less than 0.325

(B) At least 0.325, but less than 0.375

(C) At least 0.375, but less than 0.425

(D) At least 0.425, but less than 0.475

(E) At least 0.475

©ACTEX 2014

5. Fall 2000 #6 SOA

An insurance company has two insurance portfolios. Claims in Portfolio *P* occur in accordance with a Poisson process with mean 3 per year. Claims in Portfolio *Q* occur in accordance with a Poisson process with mean 5 per year. The two processes are independent.

Calculate the probability that 3 claims per year occur in Portfolio *P* before 3 claims occur in Portfolio *Q*.

(A) 0.28

(B) 0.33

(C) 0.38

(D) 0.43

(E) 0.48

6. Fall 2004 #8 SOA

For a tyrannosaur with a taste for scientists:

(i) The number of scientists eaten has a binomial distribution with $q = 0.6$ and $m = 8$.

(ii) The number of calories of a scientist is uniformly distributed on (7000, 9000).

(iii) The numbers of calories of scientists eaten are independent, and are independent of the number of scientists eaten.

Calculate the probability that two or more scientists are eaten and exactly two of those eaten have at least 8000 calories each.

(A) 0.23

(B) 0.25

(C) 0.27

(D) 0.30

(E) 0.35

©ACTEX 2014

SOLUTIONS TO PROBLEM SET 1

1. (Poisson decomposition)
The number of coins found in an hour has a Poisson distributions with $\lambda = 60(0.5) = 30$ coins per hour. This can be decomposed into 3 Poisson distributions, one for each denomination of coin:

(1) for coins of value 1 the parameter is $0.6(\lambda) = 18$
(2) for coins of value 5 the parameter is $0.2(\lambda) = 6$
(3) for coins of value 10 the parameter is $0.2(\lambda) = 6$

Since he actually finds 10 coins worth 5 cents each the expected value of all coins found is

$10(5) + 1$ (expected number of coins worth 1)
$+10$ (expected number of coins worth 10)
$= 50 + 1(18) + 10(6) = 128$

Answer C

2. (This problem was solved using trees in Example (2.103). The solution here is based purely on formulas without trees. If you understood (2.103), you do not need to read this again.)

Let *B* and *G* denote bad and good drivers respectively, and *N* denote the number of claims in a year. $P(N = n) = \frac{\lambda^n e^{-\lambda}}{n!}$.

For good drivers, $\lambda = 1$ and for bad drivers $\lambda = 5$.

$P[B] = P[G] = 0.5$
$P[N = 3 \mid G] = \frac{e^{-1}}{3!} = 0.061313$
$P[N = 3 \mid B] = \frac{5^3 e^{-5}}{3!} = 0.140374$
$P[G \cap (N = 3)] = P[G] \cdot P[N = 3 \mid G] = (0.5)(0.061313) = 0.030657$
$P[B \cap (N = 3)] = P[B] \cdot P[N = 3 \mid B] = (0.5)(0.140374) = 0.070187$
$P[N = 3] = 0.030657 + 0.070187 = 0.10084$
$P[G \mid N = 3] = \frac{P[G \cap (N = 3)]}{P[N = 3]} = \frac{0.030657}{0.10084} = 0.304$

Answer A

©ACTEX 2014

5. Total claims for company follow a Poisson distribution with $\lambda = 3 + 5 = 8$ per year. The probability that a given claim is from portfolio *P* is $\frac{3}{8}$ and from portfolio *Q* is $\frac{5}{8}$. In order that 3 claims from *P* occur before 3 claims from *Q*, the first 5 claims could have at most 2 from *Q*. This can be viewed as a binomial process with $n = 5$ and $p = \frac{3}{8}$ (the probability of *P*). If *k* is the number of claims in *P* we want

$$P[k \geq 3] = \binom{5}{3}\left(\frac{3}{8}\right)^3\left(\frac{5}{8}\right)^2 + \binom{5}{4}\left(\frac{3}{8}\right)^4\left(\frac{5}{8}\right) + \binom{5}{5}\left(\frac{3}{8}\right)^5 = 0.2752$$

Answer A

6. (Compound counting distribution)
The number *N* of scientists eaten in a day has a binomial distribution with $n = 8$ and $p = 0.6$.

$$P(N = n) = \binom{8}{n}(0.6)^n(0.4)^{8-n}$$

Since the number of calories of a scientist is uniformly distributed over (7,000, 9,000) the probability that a scientist has at least 8,000 calories is 0.5.If *k* is the number of scientists eaten who have over 8,000 calories,

$$P[k = 2 \mid N = n \geq 2] = \binom{n}{2}(0.5)^n.$$

So we need $\sum_{n=2}^{8} P[N = n]\,P[k = 2 \mid N = n]$

Grinding this out we get:

$$n = 2 \quad \binom{8}{2}(0.6)^2(0.4)^6\binom{2}{2}(0.5)^2 = 0.01032$$

$$n = 3 \quad \binom{8}{3}(0.6)^3(0.4)^5\binom{3}{2}(0.5)^3 = 0.04645$$

$$n = 4 \quad \binom{8}{4}(0.6)^4(0.4)^4\binom{4}{2}(0.5)^4 = 0.08709$$

$$n = 5 \quad \binom{8}{5}(0.6)^5(0.4)^3\binom{5}{2}(0.5)^5 = 0.08709$$

$$n = 6 \quad \binom{8}{6}(0.6)^6(0.4)^2\binom{6}{2}(0.5)^6 = 0.04899$$

$$n = 7 \quad \binom{8}{7}(0.6)^7(0.4)\binom{7}{2}(0.5)^7 = 0.01470$$

$$n = 8 \quad (0.6)^8\binom{8}{2}(0.5)^8 = 0.00184$$

Summing these we get 0.29648

Answer D

©ACTEX 2014

31. Let K be the number of customers who buy a policy. We want the probability $P(K \geq 56)$.
With the integer correction this becomes $P(K \geq 55.5)$.
To apply the central limit theorem we need $E(K)$ and $Var(K)$.
K has a binomial distribution with $n = 500$ and $q = .10$ (success probability).
Therefore, $E(K) = 500(.1) = 50$ and $Var(K) = 500(.1)(.9) = 45$.
Then,
$P(K \geq 55.5) = P(\frac{K-50}{\sqrt{45}} \geq \frac{55.5-50}{\sqrt{45}}) = P(Z \geq .82) = 1 - \Phi(.82) = 1 - .7939 = .2061.$
Answer: B

32. I. False. $E(\widehat{\theta}_n) = E(5,000n/(n+1)) = 5000n/(n+1) \neq 5000 = \theta$.
II. True. $|\theta - \widehat{\theta}_n| = |5000 - 5000n/(n+1)| = 5000/(n+1) \rightarrow 0$ as $n \rightarrow \infty$,
so that if $\delta > 0$, then $P(|\theta - \widehat{\theta}_n| < \delta) = P(\frac{5000}{n+1} < \delta) = 1$ when n is large.
Therefore, $\lim_{n\to\infty} P(|\theta - \widehat{\theta}_n| < \delta) = 1$ for any $\delta > 0$, which means that the estimator is consistent.
III. True. $MSE(\widehat{\theta}_{10}) = E[(\widehat{\theta}_{10} - \theta)^2] = E[(\frac{5000(10)}{11} - 5000)^2] = 206,612$.
Answer: D

33. If X has an inverse exponential distribution parameter θ, then $Y = \frac{1}{X}$ has an exponential distribution with mean $\frac{1}{\theta}$. We transform the given inverse exponential data set of sample x's to sample y's with $y = \frac{1}{x}$. So $\frac{1}{8000}$, $\frac{1}{10,000}$, $\frac{1}{12,000}$, $\frac{1}{15,000}$ is a sample from an exponential distribution with mean $\frac{1}{\theta}$. The mle of $\frac{1}{\theta}$ is
$\frac{\frac{1}{8000}+\frac{1}{10,000}+\frac{1}{12,000}+\frac{1}{15,000}}{4} = .00009375$.
The mle of θ is $\frac{1}{.00009375} = 10,667$. Answer: C

34. The log of the pdf is $ln\, f(x;\theta) = ln\,\theta + (\theta - 1)\, ln\, x$
and the derivative of the log of the pdf is $\frac{d}{d\theta} f(x;\theta) = \frac{1}{\theta} + ln\, x$
so the derivative of the loglikelihood is $\frac{d}{d\theta} ln\, L = \frac{n}{\theta} + \Sigma\, ln\, x$.

With the $n = 5$ values, setting the loglikelihood equal to 0 results in
$\frac{d}{d\theta} ln\, L = \frac{5}{\theta} + (ln.25 + ln.50 + ln.40 + ln.80 + ln.65) = 0$.

The mle of θ is 1.37 . Answer: E

© ACTEX 2014

Example ST-12: You are given the negative loglikelihoods associated with five models:

Model	Number of Parameters	Negative Loglikelihood
Generalized Pareto	3	219.1
Burr	3	219.2
Pareto	2	221.2
Lognormal	2	221.4
Inverse Exponential	1	224.2

Using the Likelihood Ratio Test at a 5% level of significance, determine which models are less preferable than the Generalized Pareto.

Solution: The test statistic is $2(ln\, L_A - ln\, L_B)$. For the 5% χ^2 test, the critical values are 3.84 with 1 degree of freedom and 5.99 with 2 degrees of freedom.
The generalized Pareto is automatically preferable to the Burr since it has the same number of parameters but has a higher loglikelihood value.

With model A being the generalized Pareto, the test statistic values are

Model B	Test Statistic	Degrees of Freedom
Pareto	4.2	1
Lognormal	4.6	1
Inverse Exponential	10.2	2

The likelihood ratio test prefers the Generalized Pareto to all other models, since the test statistic is greater than the critical value in all cases. We have performed 3 separate hypothesis tests. The null hypothesis for the first test is

$H_0 : X$ has a Pareto distribution and the alternative is $H_1 : X$ has a Generalized Pareto.

This is a chi-square test with $3 - 2 = 1$ degree of freedom since Generalized Pareto has 3 parameters and Pareto has 2. The other two tests have null hypothesis of lognormal and null hypothesis of inverse exponential, but the alternative is still Generalized Pareto. □

© ACTEX 2014

Chi-square goodness-of-fit test: This test is based on grouped data. The data is separated into k groupings. The groupings may be intervals $(0, c_1], \ldots, (c_{k-1}, c_k]$ (c_k may be ∞), with n_j data points in interval $(c_{j-1}, c_j]$, and a total of n data points. The groupings may be integer categories. For instance, if we are counting number of claims per day, we may have n_0 days with 0 claims, n_1 days with 1 claim, etc. We are trying to determine if a particular model distribution (perhaps a distribution that we estimated) fits the data well. For each category, we determine the probability for that category based on the model distribution, the j-th category will have model probability $\hat{p}_j$. For instance, if the random variable is the size of a loss, and if the category is the interval $(c_{j-1}, c_j]$, then $\hat{p}_j$ is the probability that the random variable is in the interval; this probability is $\hat{p}_j = F^*(c_j) - F^*(c_{j-1})$. If the random variable is the number of claims in a day, and if the category is the integer j, then $\hat{p}_j$ is the probability that there will be j claims in one day.

If the total number of observations is n, then according to the model distribution, the expected number of observations in category j will be $E_j = n\hat{p}_j$. The observed number of data points in category j is denoted is n_j (also denoted O_j in the Loss Models book). In the Loss Models book, O_j is also written as np_{n_j}.

The chi-square goodness-of-fit statistic is

$$\boldsymbol{Q = \sum_{j=1}^{k} \frac{(n_j - E_j)^2}{E_j} = \sum_{j=1}^{k} \frac{n^2(p_{nj} - \hat{p}_j)^2}{E_j} = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j}}$$

Q has an approximate chi-square distribution with $k - r - 1$ degrees of freedom, where r is the number of parameters that have been estimated in the model distribution, and k is the number of categories. The null hypothesis is still the hypothesis that the model distribution fits the data well. The null hypothesis is rejected at the 5% level of significance if $Q > \chi^2_{k-r-1,.05}$ (the 95-th percentile of the chi-square distribution with $k - r - 1$ degrees of freedom; any level of significance can be tested). Rejection of the null hypothesis is an indication that the model does not fit the data well.

© ACTEX 2014

PRACTICE TEST 7

1. A normal population is known to have a mean of 10 and unknown variance of $\sigma^2 > 0$. Based on a random sample of size n, what is the maximum likelihood estimate of σ^2?

A) $\frac{1}{n-1}\Sigma(x_i-\overline{x})^2$

B) $\frac{1}{n}\Sigma(x_i-\overline{x})^2$

C) $\frac{1}{n+1}\Sigma(x_i-\overline{x})^2$

D) $\frac{1}{n-1}\Sigma(x_i-10)^2$

E) $\frac{1}{n}\Sigma(x_i-10)^2$

2. Consider the regression model: $y_i = \beta_1 x_i + \beta_2 w_i + \epsilon_i$, where $i = 1,\ldots,n$ and $(x_1,w_1),\ldots,(x_n,w_n)$ are fixed constants. Let the independent random variables $\epsilon_1,\ldots,\epsilon_n$ have mean 0 and common variance σ^2. What is the least squares estimate of β_1 if it is given that $\sum_{i=1}^{n} x_i^2 = 1$ and $\sum_{i=1}^{n} w_i^2 = 1$?

A) $\frac{\Sigma(x_i-\overline{x})(y_i-\overline{y})}{\Sigma(x_i-\overline{x})^2}$

B) $\frac{\Sigma x_i y_i}{\Sigma x_i^2}$

C) $\frac{\Sigma y_i-\beta_2\Sigma w_i}{\Sigma x_i}$

D) $\frac{(\Sigma y_i w_i)(\Sigma w_i x_i)-\Sigma y_i x_i}{(\Sigma w_i x_i)^2-1}$

E) $\frac{(\Sigma y_i x_i)(\Sigma w_i x_i)-\Sigma y_i w_i}{(\Sigma w_i x_i)^2-1}$

3. Let $X_1,\ldots,X_n$ be a random sample of size n from a distribution with density function $f(x) = \frac{2x}{\theta^2}$ for $0 \le x \le \theta$. What is the maximum likelihood estimator of θ?

A) $max(X_1,\ldots,X_n)$

B) $min(X_1,\ldots,X_n)$

C) $(X_1,\ldots,X_n)^{1/n}$

D) $\bar{X}$

E) $\frac{3\bar{X}}{2}$

© ACTEX 2014

PRACTICE TEST 7 SOLUTIONS

1. Given a density function $f(x;\theta)$ that depends upon the value of a parameter θ, and given a random sample of values $x_1,\ldots,x_n$, the maximum likelihood estimate of θ is that value of θ which maximizes the likelihood function
$L(\theta) = L(\theta; x_1,\ldots,x_n) = f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$. In this problem, we are dealing with a normal population with mean 10 and unknown variance σ^2 (parameter σ^2). Therefore, $f(x;\sigma^2) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot exp\left(-\frac{(x-10)^2}{2\sigma^2}\right)$.
Then $L(\sigma^2)$ is equal to
$\prod_{i=1}^{n} f(x_i;\sigma^2) = \prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot exp\left(-\frac{(x_i-10)^2}{2\sigma^2}\right)\right] = \frac{1}{(2\pi)^{n/2}\cdot\sigma^n}\cdot exp\left(-\sum_{i=1}^{n}\frac{(x_i-10)^2}{2\sigma^2}\right)$.

Maximizing $L(\sigma^2)$ is equivalent to maximizing its natural logarithm:
$g(\sigma) = \ln L(\sigma^2) = \ln\prod_{i=1}^{n} f(x_i;\sigma^2)$
$= \sum_{i=1}^{n}\ln f(x_i;\sigma^2) = -\frac{n}{2}\cdot\ln(2\pi) - n\cdot\ln(\sigma) - \sum_{i=1}^{n}\frac{(x_i-10)^2}{2\sigma^2}$.

Setting $\frac{d}{d\sigma}g(\sigma) = 0$ and solving for σ results in $-\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(x_i-10)^2}{\sigma^3} = 0$, or equivalently, $\sigma^2 = \frac{\Sigma(x_i-10)^2}{n}$. The second derivative test shows that
$\frac{d^2}{d\sigma^2}g(\sigma) = \frac{n}{\sigma^2} - 3\cdot\sum_{i=1}^{n}\frac{(x_i-10)^2}{\sigma^4}$, so that if $\sigma^2 = \frac{\Sigma(x_i-10)^2}{n}$, then
$\frac{d^2}{d\sigma^2}g(\sigma) = \frac{n^2}{\Sigma(x_i-10)^2} - \frac{3n^2}{\Sigma(x_i-10)^2} < 0$.

Therefore, $g(\sigma)$ (and hence $L(\sigma^2)$) has maximum at $\sigma^2 = \frac{\Sigma(x_i-10)^2}{n}$. This is the maximum likelihood estimator. Note that in general it is not necessary to apply the second derivative test; because of the nature of likelihood functions, the critical point of $L(\theta)$ (or its natural log) will be the point at which the likelihood is maximized. Answer: E

2. The least squares estimates for β_1 and β_2 in the given model are found by minimizing $SS = \sum_{i=1}^{n}(y_1-\beta_1 x_i-\beta_1 w_i)^2$ with respect to β_1 and β_2. This can be done by taking the partial derivative of SS with respect to β_1 and setting it equal to zero, then doing the same for β_2, and then solving for β_1 and β_2 from the resulting equations:
$\frac{\partial}{\partial\beta_1}SS = 2[\sum(y_i-\beta_1 x_i-\beta_2 w_i)(-x_i)] = 0$, $\frac{\partial}{\partial\beta_2}SS = 2[\sum(y_i-\beta_1 x_i-\beta_2 w_i)(-x_i)] = 0$
or equivalently, $\beta_1\cdot\sum x_i^2+\beta_2\cdot\sum x_i w_i = \sum y_i x_i$ and $\beta_1\cdot\sum x_i w_i + \beta_2\cdot\sum w_i^2 = \sum y_i w_i$.

© ACTEX 2014