Tutorial "Performance Evaluation Techniques" Second Problem Sheet

Dr.-Ing. Andreas Willig
Hasso-Plattner-Institute, University of Potsdam
email: willig@hpi.uni-potsdam.de

April 29, 2003

Date of assignment	Apr. 30, 2003 (in the tutorial)
Date of submission	May 14, 2003 (in the tutorial)

Problem 1:
For $t \in \mathbb{R}^{\geq 0}$ be X_t a random variable having normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$ (this is written as $X_t \sim N(0,1)$) and furthermore all X_t are independent. We define the stochastic process $(Y_t)_{t\in\mathbb{R}^{\geq 0}}$ by:

$$Y_t = t + X_t$$

- Is $(Y_t)_{t\in\mathbb{R}^{\geq 0}}$ stationary?
- Has $(Y_t)_{t\in\mathbb{R}^{\geq 0}}$ independent increments?
- Has $(Y_t)_{t\in\mathbb{R}^{\geq 0}}$ stationary increments?
- Find $E[Y_t]$ and $\text{Var}[Y_t]$

	2

Problem 2:
Be $(N(t))_{t\in\mathbb{R}^{\geq 0}}$ a Poisson Process with rate λ. We are given that one arrival occured in the interval $(0,t]$ (hence $N(t) = 1$) but we do not know when this arrival happened. Let Y denote the random variable of the first arrival time, the range of Y is $(0,t]$. Show that Y has a uniform distribution (hence, its distribution function is given by $F(y) = \Pr[Y \leq y] = \frac{y}{t}$). Hint: compute $\Pr[Y \leq y \,|N(t) = 1]$.

	2

Problem 3:
Be $(N(t))_{t\in\mathbb{R}^{\geq 0}}$ a Poisson Process with rate λ, and Y a continuous non-negative random variable independent of $(N(t))_{t\in\mathbb{R}^{\geq 0}}$ with density function $f(y)$.

- Show that $G_{N(Y)}(z) = E\left[z^{N(Y)}\right]$ can be expressed as:

  $$G_{N(Y)}(z) = \mathcal{M}_Y(-\lambda(1-z))$$

  (i.e. the z-Transform of $N(Y)$ can be expressed in terms of the moment generating function for Y). Hint: law of total probability for continuous random variables.

- Use $G_{N(Y)}(z)$ to find $E[N(Y)]$ and $\mathrm{Var}[N(Y)]$

	4

Problem 4:
Sometimes one is interested in making predictions, e.g. we know that a file transfer already took t time units, what is the probability that it takes another $d > 0$ time units?
A random variable X is called *heavy-tailed*, when

$$\Pr[X > t] \sim t^{-\alpha} \qquad (t \to \infty)$$

for $0 < \alpha < 2$. (A function $f(x)$ behaves asymptotically as $g(x)$, written as $f(x) \sim g(x)$, if there exists some $c \in \mathbb{R}, c \neq 0$ such that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = c$ holds [$g(x) \neq 0$ required]).

- Show that the exponential distribution with parameter λ is not heavy-tailed
- Compute $\lim_{t\to\infty} \Pr[X > t + d \,|X > t]$ for the exponential distribution
- Compute $\lim_{t\to\infty} \Pr[X > t + d \,|X > t]$ for a heavy-tailed distribution
- Compare and interpret the results

In many studies it has been found that the distribution of file sizes in a UNIX file system and the document sizes of a web server are heavy-tailed.

	3

Bonus Problem 5:
The distribution and density of a random variable X with Pareto distribution are given by:

$$\begin{aligned} F(x) &= \Pr[X \le x] = 1 - \left(\frac{k}{x}\right)^{\alpha} \\ f(x) &= \alpha \cdot k^{\alpha} \cdot x^{-\alpha-1} \end{aligned}$$

for $\alpha > 0$, $k > 0$ and $x \ge k$. For $0 < \alpha < 2$ the Pareto distribution is heavy-tailed. Be X such a Pareto random variable with parameters $\alpha = 1$, $k = 1$. Furthermore, be Y an exponential random variable with parameter $\lambda = 1$.
Plot both $\Pr[X > x]$ and $\Pr[Y > x]$ for $1 \le x \le 100$ using a doubly-logarithmic scale (the `gnuplot` program has the `set logscale x` and `set logscale y` commands for this). In this kind of plots heavy-tailed variables have in general a linear shape.

	2