Lecture 26: Mar 9, Sum of a random number of random variables

26.1 The expectation (Ross P.369)

Let X_i $i = 1, 2, ...$ all have mean μ.

Let N be a random integer, with N independent of the X_i;

we are interested in $T \;=\; \sum i = 1^N X_i$.

Example: X_i is weight of person; N is number of people entering elevator; T is total weight.

Or; X_i is money spent by person i; N is number of people in store; T is total intake.

$$\mathrm{E}(\sum_{i=1}^{N} X_i \mid N = n) \;=\; \mathrm{E}(\sum_{i=1}^{n} X_i) \;=\; \sum_{i=1}^{n} \mathrm{E}(X_i) \;=\; n\mu$$

$$\mathrm{E}(\sum_{i=1}^{N} X_i) \;=\; \mathrm{E}(\mathrm{E}(\sum_{i=1}^{N} X_i \mid N)) \;=\; \mathrm{E}(N\mu) \;=\; \mathrm{E}(N)\mathrm{E}(X)$$

Note we do use the independence of N and X_i; $\mathrm{E}(X_i)$ is unchanged by fixing $N = n$.

26.2 The variance (Ross P.382)

Let X_i $i = 1, 2, ...$ be (pairwise) independent, all with mean μ and variance σ^2.

Let N be a random integer, with N independent of the X_i.

We are interested in $T \;=\; \sum i = 1^N X_i$; examples as above.

$$\mathrm{var}(\sum_{i=1}^{N} X_i \mid N = n) \;=\; \mathrm{var}(\sum_{i=1}^{n} X_i) \;=\; \sum_{i=1}^{n} \mathrm{var}(X_i) \;=\; n\sigma^2$$

$$\mathrm{var}(\sum_{i=1}^{N} X_i \mid N) \;=\; N\sigma^2 \quad \text{and} \quad \mathrm{E}(\sum_{i=1}^{N} X_i \mid N) \;=\; N\mu$$

$$\mathrm{var}(\sum_{i=1}^{N} X_i) \;=\; \mathrm{E}(\mathrm{var}(\sum_{i=1}^{N} X_i \mid N)) \;+\; \mathrm{var}(\mathrm{E}(\sum_{i=1}^{N} X_i \mid N))$$

$$= \; \mathrm{E}(\sigma^2 N) \;+\; \mathrm{var}(\mu N) \;\; \sigma^2 \mathrm{E}(N) \;+\; \mu^2 \mathrm{var}(N)$$

26.3 Examples

(i) People entering an elevator have mean weight 160lb, with variance 400lb^2. The number of people, N entering is Poisson with mean 4. What are the mean and variance of the total weight, T.

$\mathrm{E}(T) \;=\; \mathrm{E}(N) \times 160 \;=\; 640$lb. $\qquad \mathrm{var}(T) \;=\; 400 \times \mathrm{E}(N) \;+\; 160^2 \times \mathrm{var}(N) \;=\; 104000$lb^2 (st.dev 322 lb).

(ii) A coin with probability of heads p, is tossed N times, where N is Poisson with mean (and variance) μ. What are the mean and variance of the number of heads, T.

Given $n = N$, $X_i \;\sim\; Bin(1, p)$, $T = \sum_i X_i \;\sim\; Bin(n, p)$. $\mathrm{E}(X_i) \;=\; p$, $\mathrm{var}(X_i) \;=\; p(1-p)$.

$T = \sum_1^N X_i$, $\mathrm{E}(T) \;=\; \mu p$, $\mathrm{var}(T) \;=\; \mathrm{var}(N)p^2 + \mathrm{E}(N)p(1-p) \;=\; \mu p^2 + \mu p(1-p) \;=\; \mu p$.

Lecture 27: Mar 11, More Conditional Expectations; using mgf's

27.1 Ch 7; Exx 56

A number X of people enter an elevator at the ground floor; $X \sim \mathcal{P}o(10)$.

There are n upper floors and each person (independently) gets off at floor k with probability $1/n$. Find the expected number of stops.

Probability no-one gets off at a particular floor is $(1-1/n)^X$. So expected number of floors the elevator does **not** stop is $\mathrm{E}(((n-1)/n)^X) = \exp(10((n-1)/n)-1) = \exp(-10/n)$.

So expected number of stops is $n(1-\exp(-10/n))$.

27.2 Binomial/Poisson hierarchy

We saw if X, Y are independent Poisson, then $X|(X+Y)$ is Binomial.

We saw if $(X|N)$ Binomial, and N Poisson, then overall X has mean equal to variance (like a Poisson), so ....

If $X \sim Bin(np)$, $m_X(t) = \mathrm{E}(e^{tX}) = (q+pe^t)^n$ where $q = 1-p$.

If $Y \sim \mathcal{P}o(\mu)$, $m_Y(t) = \mathrm{E}(e^{tY}) \equiv \mathrm{E}(z^Y) = \exp(\mu(z-1))$, where $z \equiv e^t$.

Now if $X|Y \sim Bin(Y,p)$, and $Y \sim \mathcal{P}o(\mu)$,

$$m_X(t) = \mathrm{E}(e^{Xt}) = \mathrm{E}(\mathrm{E}(e^{Xt} \mid Y) = \; = \mathrm{E}((q+pe^t)^Y) = \exp(\mu((q+pe^t)-1) = \exp(\mu p(e^t-1)).$$

So by uniqueness of mgf, X is Poisson with mean μp.

27.3 Poisson/Gamma hierarchy gives Negative Binomial

If $Y \sim \mathcal{P}o(\mu)$, $m_Y(t) = \exp(\mu(e^t-1))$. If $Z \sim G(r,\lambda)$, $m_Z(t) = (\lambda/(\lambda-t))^r$.

If $Y|Z \sim \mathcal{P}o(Z)$, and $Z \sim G(r,\lambda)$.

$$m_Y(t) = \mathrm{E}(e^{Yt}) = \mathrm{E}(\mathrm{E}(e^{Yt} \mid Z) = \mathrm{E}(\exp(Z(e^t-1))) =$$
$$m_Z(e^t-1)) = (\lambda/(\lambda-(e^t-1)))^r = (p/(1-qe^t))^r,$$

where $p = \lambda/(\lambda+1)$, $q = 1-p = 1/(\lambda+1)$. But this is e^{-tr} times the mgf of a NegBin (r,p).

i.e. it is the NegBin where we count the failures before the rth success, and not the r successes.

So, by uniqueness of mgf, this is the marginal pmf of Y.

27.4 Sum of a Geometric number of independent Exponentials

If $X_i \sim \mathcal{E}(\lambda)$; $m_{X_i}(t) = \lambda/(\lambda-t)$.

If $Y \sim Geo(p)$, $m_Y(t) = \mathrm{E}(z^Y) = pz/(1-qz)$, where $z \equiv e^t$.

Let $W = \sum_1^Y X_i$. Given $Y = n$, $m_W(t) = \prod m_{X_i}(t) = (m_X(t))^n$.

Then $m_W(t) = \mathrm{E}(e^{Wt}) = \mathrm{E}(\mathrm{E}(e^{Wt} \mid Y)) = \mathrm{E}(m_X(t)^Y) = pm_X(t)/(1-qm_X(t))$

$$= p\lambda/(\lambda-t-q\lambda) = p\lambda/(p\lambda-t).$$

But this is the mgf of an exponential $\mathcal{E}(p\lambda)$, so by uniqueness of mgf, $W \sim \mathcal{E}(p\lambda)$.

This makes sense; the exponential distribution has the forgetting property. The geometric distribution has the forgetting property. So summing a "forgetting property" number of "forgetting property" random variables, should give us the "forgetting property" pdf back again.

Note if we sum a fixed number n of independent exponentials, $\mathcal{E}(\lambda)$, we get a $G(n,\lambda)$, so this example is equivalent to $X|Y \sim G(Y,\lambda)$, and $Y \sim Geo(p)$.